UNIVERSITY OF LONDON THESIS

Degree PhD Year 2007 Name of Author GARETH JAMES MANNALL

Characterisation of the effect of process factors upon protein refolding yield

A thesis submitted to the University of London for the degree of
DOCTOR OF PHILOSOPHY
By

Gareth James Mannall BSc (Hons), MSc
University College London
2007

Department of Biochemical Engineering
University College London,
Torrington Place,London,
WC1E 7JE

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Declaration

I, Gareth James Mannall, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Acknowledgements

I wish to sincerely thank Professor Nigel J Titchener Hooker and Dr Paul Dalby for the invaluable supervision they have given me throughout the course of this study. Their advice and support has been incredibly useful.

I would like to acknowledge the financial support provided by the Engineering and Physical Sciences Research Council and the IMRC for Bioprocessing without whom this study would not have been possible.

I would also like acknowledge Professor Anton Middelberg and Professor Howard Chase for their advice during this project.

I owe a great debt to all the staff in the Department of Biochemical Engineering, whom have assisted and tolerated me during my time at University College London.

I would also like to thank my friends and colleagues in the department, for their advice and friendship during my time at UCL.

Finally I would like to thank my parents whom have provided encouragement, advice and support during this study and my former years.

Abstract

Expression of recombinant protein in bacteria such as *Escherichia coli* can result in the formation of inclusion bodies. Active protein is derived from inclusion bodies by protein refolding. Refolding yields are often poor, and are a bottleneck in such recombinant protein processes. Work has addressed refolding process issues, but several areas remain poorly understood. This thesis aims to understand the importance of process parameters upon refolding yields.

Previous studies highlighted that mixing can affect refolding, but have failed to establish why. Use of a two impeller system with operation of small paddle impeller (Re= 2000) in the proximity of the injection point, revealed the importance of energy dissipation experienced by injected denatured protein upon lysozyme refolding yields.

A factorial experiment studying the effect of factors on lysozyme refolding yield in fed-batch revealed effects and interactions between physical and chemical process parameters. GdHCl concentration (1.2M) and redox ratio (2:1 red:ox) had the greatest effect. A graphical (windows of operation) approach revealed high GdHCl concentration (1.2 M) and redox ratios at or above unity gave greatest yields in the minimum time.

Industry typically refolds from impure inclusion bodies. A series of studies using trypsinogen IB detailed the effect of process contaminants and the efficacy of steps to remove them, upon refolding. Analysis of the effect of centrifugation conditions on IB purity demonstrated the compromise between the levels of IB purity and recovery achieved, together with the removal of key contaminants and the refolding yield obtained, with an optimum of $Q/\Sigma = 23.08 \times 10^{-9}$ m/s. A second study looked at IB washing and revealed that maximising the area for washing is key to IB purity but not to refolding yield. Use of Triton X was key to maximising refolding yields.

Table of Contents

List of Figures

List of Tables

Nomenclature

α	Micromixidness ratio (-)
ε	Energy dissipation rate (Wkg^{-1})
μ	Viscosity (Pa.s)
ρ	Density (kgm^{-3})
Σ	Sigma factor (m^2)
ν	Kinematic viscosity (m^2s^{-1})
APS	Ammonium persulfate
Arg	Arginine
Asp	Aspartate
BSA	Bovine serum albumin
D_D	Disc diameter (mm)
D_I	Impeller diameter (mm)
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
DNAse	Deoxyribonuclease I
D_T	Tank diameter (mm)
DTT	Dithiothreitol
E.coli	Escherichia coli
EBA	Expanded bed (chromatography)
EDTA	Ethylenediaminetetraacetic acid
G	Guanidine Hydrochloride concentration (M)
GdHCl	Guanidine Hydrochloride
H_B	Blade height (mm)
HCl	Hydrochloride
HEWL	Hen egg white lysozyme

HIC	Hydrophobic interaction chromatography
His	Histidine
Homog	Homogenization buffer
H_T	Tank height (mm)
I	Intermediate
IB	Inclusion bodies
IEX	Ion exchange (chromatography)
IMAC	Immobilized metal affinity chromatography
Ka	Acid dissosciation constant
L-BAPNA	N-α-benzoyl-L-arginine 4-nitroanilide
LDH	Lactate dehyrogenase
M	Mol/L
min	Minute(s)
N	Native protein (refolding mechanisms)
N	Impeller speed (s^{-1})
n	Mol number
NTSB	2-nitro-5-(thiosulfo)-benzoate
OD	Optical density
Omp	Outer membrane protein
Ox	Oxidised form of redox ratio (disulfide form)
Pa	Pascals
PEG	Polyethylene glycol
Q	Flow rate (mL/ min)
R	Redox ratio (reduced:oxidised) (-)
Re	Reynolds number (-)
Re_b	Bulk impeller Reynolds number (-)
Red	Reduced form of redox reagent (free thiol form)
Re_m	Mini impeller Reynolds number (-)

RNA	Ribonucleic acid
RP-HPLC	Reverse Phase High Performance Liquid Chromatography
RR	Redox ratio
RSD	Residual standard deviation
SDS-PAGE-	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
Ser	Serine
TB	Terrific broth
TE	Tris EDTA (buffer)
TEMED	N',N',N',N'-tetramethylethyldiamine
TFA	Trifluoroacetic acid
t_{inj}	Time for injection (s)
t_m	Mixing time (s)
Tris	Tris Hydroxymethylaminoethane
Triton X	Octyl phenol ethoxylate
U	Unfolded protein
V	Volume
W_B	Blade width (m)
W_B	Blade width (m)
x_s	Segregation Index (-)

1 Literature Review

1.1 Introduction

The existence of diseases caused by genetic or environmentally-induced deficiencies in essential proteins provide the need for replacement therapy (Pain, 1987; Russell and Clarke, 1999). Replacement therapy has been useful for diseases such as diabetes and dwarfism, where treatment with insulin and growth hormone respectively attenuate the associated debilitating effects. Such proteins were originally sourced from animals or cadavers, however, these sources run the risk of contamination, especially that of a viral nature (Pain, 1987). The advent of recombinant DNA technology (Cohen et al., 1973; Jackson et al., 1972; Lobban and Kaiser, 1973; Mertz and Davis, 1972) allowed expression of required proteins in host cells such as *Escherichia coli*. Expression in such systems can often result in the aggregation of recombinant protein as insoluble inclusion bodies. Active protein is derived from inclusion bodies by a process called refolding. Refolding often presents a bottleneck in recombinant protein processes. This thesis details a series of studies not only to understand the effect of factors within refolding process, but also the influence of upstream processing steps. The following literature review aims to provide background to the areas of research in this thesis. Providing an introduction to IBs and their processing, and a detailed analysis of refolding.

1.2 Inclusion Bodies

1.2.1 Structural Characteristics and Composition of Inclusion Bodies

Inclusion bodies (IBs) are dense, refractile, protease resistant aggregates of mis-folded protein, commonly formed when over-expressing recombinant proteins in bacteria such as *E.coli* (Carrió and Villaverde, 2002). Inclusion bodies were first observed on over expression of β-galactosidase (Cheng et al., 1981). The dense refractile nature of these aggregates means that they can be easily visualised by electron microscopy (Figure 1.1) and are often observed by light microscopy (Fahnert et al., 2004). Misfolded protein within the IB is thought to be stabilised by hydrophobic interactions (Rudolph, 1996). Evidence suggests that proteins within the IB may be partially structured (Mitraki and King, 1989; Oberg et al., 1994; Przybycien et al., 1994). There is an apparent increase in the amount of beta-pleated sheets and a decrease in the number of alpha helices when IB and native protein structure are compared (Oberg et al., 1994). Evidence suggests that IBs may form from native protein, but this appears to be exceptional (Tokatlidis et al., 1991).

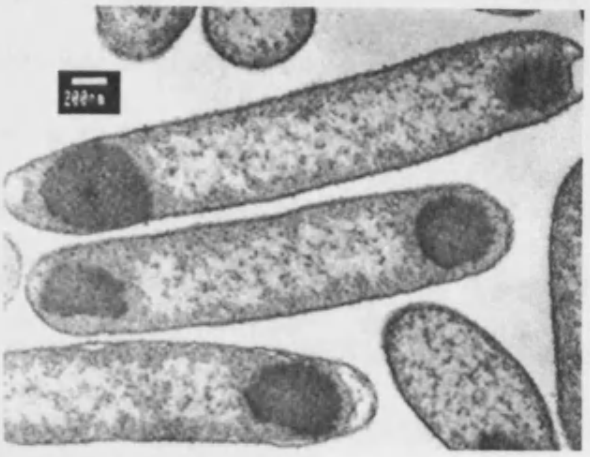


Figure 1.1: Electron microscopy of Inclusion body containing cells from (Betts, 1998)

The primary component of IBs are misfolded recombinant product (Mattes, 2006; Rinas and Bailey, 1992; Rinas and Bailey, 1993). IB preparations are typically contaminated with peptidoglycans, membrane proteins, ribosome component proteins, lipids and nucleic acids (Fahnert et al., 2004; Thatcher, 1990). These IB contaminants result from inefficiencies in IB isolation steps (see section 1.2.4), where cell debris co-sediments with IBs. Evidence suggests that contaminant protein may also be directly incorporated into the IBs (Rinas and Bailey, 1993; Valax and Georgiou, 1993). Studies by Valax and Georgiou suggested that under various growth conditions and for certain strains used, contaminating polypeptides formed between 5% and 50% of IB mass, whilst phospholipids made up 0.5-15%, and nucleic acids formed only a minor proportion (Valax and Georgiou, 1993). A major class of proteinacious contaminant are the outer membrane proteins: OmpF, OmpC, OmpA which contaminate via co-precipitation of cell debris components (Rinas and Bailey, 1992). Other typical proteinacious contaminants include the molecular chaperones GroEL, DnaK and

the small heat shock proteins IbpA and IbpB (Carrió and Villaverde, 2002; Jürgen et al., 2000). 2D SDS-PAGE of recombinant protein from IBs did not show a single spot, but rather a smear of multiple spots in the direction of isoelectric focussing (Fahnert et al., 2004; Rinas et al., 1993). The result suggested that the structure of the recombinant protein within an IB was heterogeneous. This may have been brought about by chemical modification in downstream processing steps, the existence of the protein as different folding intermediates, or precipitation during isoelectric focussing. IBs may also contain shortened forms of the recombinant product from incorrect translation or proteolysis, and abnormally elongated versions of the product (Bowden and Georgiou, 1990; Geli et al., 1989; Hart et al., 1990).

IBs are generally thought of as aggregates of misfolded protein that are inaccessible and therefore resistant to proteolysis (Fahnert et al., 2004). Evidence suggests that a variety of different conformational states of proteins exist in IB (Bowden et al., 1991; Carrió et al., 2000), possibly suggesting conformational flexibility not only in the proteins themselves, but within the IB as a whole (Carrió and Villaverde, 2002). Further to this Carrio and Villaverde have observed refolding and simultaneous aggregation to form an IB in actively producing recombinant bacteria (Carrió and Villaverde, 2001). An IB is therefore thought to be a dynamic mass with both protein influx and flux (Carrió and Villaverde, 2002). In the absence of *de-novo* protein synthesis, IBs were observed to be broken down, with a resultant increase in both soluble protein and biological activity of the bacterium (Carrió and Villaverde, 2001). This suggests that a proportion of the IB is still in a form that may be refolded and possibly undergo proteolysis (Carrió and Villaverde, 2002). It is therefore thought that IBs

are plastic structures rather than simply just dead ends of unsuccessful refolding, where misfolded protein accumulates until chaperones and proteolytic enzymes can refold or digest them (Carrió and Villaverde, 2002).

1.2.2 Why Do IBs Form?

1.2.2.1 Competition Between Folding and Aggregation.

The formation of IBs is primarily determined by the competition between folding and aggregation, which is directly related to the rate of protein synthesis (Rudolph, 1996). Greater rates of synthesis can overwhelm the cells folding apparatus, resulting in a high concentration of misfolded protein, which is prone to aggregation. Refolding intermediates often aggregate because of their exposed hydrophobic regions (Ellis, 1997). It is thought that the folding behaviour of the protein rather than the protein's structural qualities (size, relative hydrophobicity etc) primarily determine the likelihood to aggregate (Rudolph, 1996). Despite this the formation of IB aggregates can be influenced by rate-limiting structural components such as disulfide bonds and the need for post-translational modifications in the native structure (Fahnert et al., 2004).

1.2.2.2 Structural Characteristics of Proteins Being Expressed

Structural characteristics of expressed protein can make them more susceptible to aggregation as IB. Formation of disulfide bonds in protein is a rate-limiting step for folding. The greater the number of cysteine molecules there are the greater the number of possible pairings. Time and appropriate redox conditions are required to find the correct pairings. The reducing environment of the cytoplasm in *E.coli* is not conducive to disulfide shuffling and means that disulfide containing

proteins expressed in this cellular compartment are often deposited as IBs (Fahnert et al., 2004). By switching off the expression of the proteins promoting a reducing environment (thioredoxin reductase and glutathione reductase) an environment better suited to correct disulfide bonding may become apparent and result in a reduction of IB synthesis (Bessette et al., 1999; Derman et al., 1993; Proba et al., 1995).

E.coli are incapable of producing posttranslational modifications of protein, such as glycosylation. Glycosylation can influence activity and degradation characteristics (Fahnert et al., 2004). In addition glycosylation may affect folding behaviour and solubility. Protein glycosylated in its native eukaroytic form may be prone to aggregation when expressed in *E.coli* (Fahnert et al., 2004).

1.2.2.3 Prokaryotic Versus Eukaryotic Expression Systems

Recombinant protein expressed in *E.coli* is often eukaryotic, and differences in expression systems can make soluble expression difficult. Prokaryotes such as *E.coli* translate RNA to protein more quickly and tend to fold protein after complete translation. In contrast, eukaryotes translate RNA more slowly and fold proteins both co-translationally and post-translationally (Chen et al., 1995; Ellis and Hartl, 1999; Frydman et al., 1999). The slower rate of translation in eukaryotes allows for the sequential folding of domains (Ujvari et al., 2001). The inability to perform this sequential refolding in prokarya may result in a greater degree of misfolding when expressing proteins of eukaryotic origin. High levels of expression achieved in prokaryotes increase the likelihood of aggregation, by

virtue of its dependence on the concentration of unfolded protein and total protein concentration in the cell (Ellis, 2001).

Prokaryotic chaperones are non-catalytic and are less able to promote the formation of native protein than the catalytic eukaryotic chaperones. It has been discovered that the prokaryotic chaperones GroEL and DnaK are the main antagonistic controllers of the size and number of IBs (Carrió and Villaverde, 2003). GroEL knock-down mutants produced lower amounts of IBs, but greater numbers of aggregates. DnaK knock-down mutants produced much larger IBs containing more protein than wild type cells. It was concluded that DnaK prevents IB formation by suppressing aggregation, whilst GroEL transits protein from soluble to insoluble fractions promoting aggregation (Fahnert et al., 2004).

1.2.2.4 Physiological Factors and Cultivation Conditions

The copy number of the recombinant gene directly affects the amount of product accumulated in an amplified cascade manner (Fahnert et al., 2004). Large copy numbers can enable greater synthesis rates of protein, which may overwhelm the cellular refolding apparatus and cause aggregation. In a similar manner the use of strong promoters can increase expression levels, increasing the likelihood of aggregation (Fahnert et al., 2004). As a direct consequence of this, the levels of chaperones and proteases critically affect the amount of IBs formed. Only when sufficient amounts are available can IB be broken down (Carrió and Villaverde, 2002).

Cultivation conditions have also been shown to have a major effect on the levels of protein aggregation as IBs, including high temperature, high cell densities, and high concentrations of inducer (Choi et al., 2000; Schein and Noteborn, 1988;

Wang et al., 1999). High temperatures are thought to allow the thermodynamic threshold for intermolecular reactions to be crossed, permitting aggregation. Higher levels of inducer affect the number of copies of mRNA made at a time and the amount of protein produced, increasing the likelihood of aggregation.

1.2.3 Inclusion Bodies as a Double-edged Sword

It is evident that a variety of factors influence the formation of inclusion bodies. Adaptation of such factors may prevent IB formation, but this is often difficult to achieve. The deposition of recombinant protein as misfolded aggregates, rather than soluble protein, is undesirable. Active, renatured protein can be derived from IBs, but this requires the use of stringent refolding procedures (See section 1.4.3-1.4.7) (Carrió and Villaverde, 2002). IBs therefore provide a significant obstacle for the generation of a large amount of recombinant protein in bacterial systems. Despite these apparent disadvantages, expression of proteins in IBs offers several advantages over soluble methods.

- Very high levels of expression can be achieved in inclusion bodies; levels of 30% of total cell protein or 8.5 g/L culture media have been reported (Li et al., 2004).
- Some protein products are toxic to the bacterial host, expression as insoluble IBs can increase both the viability of the cell and the yield of protein derived (Li et al., 2004).
- Inclusion bodies can be separated from other cell components more easily. Inclusion bodies are first released by mechanical or chemical means and then separated from lighter components by differential centrifugation (Middelberg, 2002).

- Inclusion bodies are a source of relatively pure polypeptide, containing typically 40-50% of desired polypeptide (Buswell et al., 2002). Once inclusion bodies have been isolated purification of the recombinant protein can be accomplished easily after refolding steps (Carrió and Villaverde, 2002).
- `Recombinant proteins are more stable in their inclusion body form than in their native state, as they are protected in the most part from proteolytic degradation (De-Bernardez Clark, 2001).
- Optimisation of the refolding of a protein is controlled by a limited number of parameters by contrast with the complexity of optimising soluble *in vivo* production (Fahnert et al., 2004).

Various proteins have been manufactured by the inclusion body route: Bovine somatotrophin (Storrs and Przybycien, 1991), tissue plasminogen activator, (Datar et al., 1993) and human insulin (Petrides et al., 1995). Details of steps that are used to produce and purify inclusion bodies ready for refolding are detailed in the next section.

1.2.4 Inclusion Body Preparation

A summary of inclusion body preparation methods is detailed in Figure 1.2.

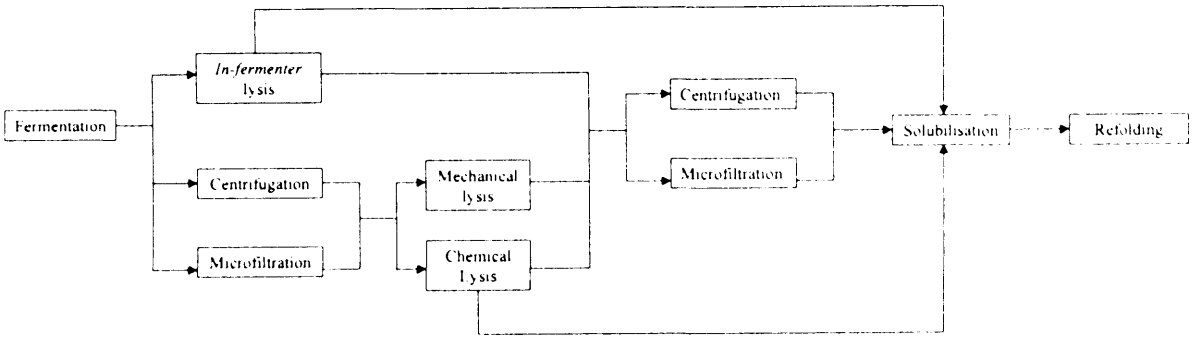


Figure 1.2: Inclusion body isolation and solubilization procedures.

1.2.4.1 Fermentation

Ideally the generation of recombinant protein as IBs should produce high quantities of high purity protein aggregates. It is apparent from earlier sections that to produce high quality IBs, high rates of the synthesis of aggregation-prone proteins is required and may be aided by low intracellular content of chaperones (especially DnaK) and proteolytic enzymes (Fahnert et al., 2004). IB-producing strains typically use strong promoters on plasmids and high concentrations of inducers are typically added to effect high synthesis rates.

1.2.4.2 Cell Lysis

Induced cells from the fermentation are harvested by centrifugation or microfiltration to remove media containing secreted proteins. Harvested cells may then be resuspended in buffer and subjected to mechanical forces to permit cell lysis and release inclusion bodies. In the laboratory this may be achieved using a French press (Bhat et al., 1991; Bird et al., 1988), where high pressures are used to shear cells by forcing them through small holes. Lab-scale lysis may be also be achieved by sonication (Geng et al., 2004; Van Kimmenade et al., 1988) where ultrasonic sound waves are used to disrupt cells walls, lysis by this method can be improved by the addition of lysozyme (Wilhelm et al., 1990; Winkler et al., 1986). At industrial scale, cells are typically disrupted by high-pressure homogenisation where cells are sheared by being forced through a valve at high pressure (typically above 500 bar) and impacting on an impact ring (Falconer et al., 1998; Jin et al., 1994; Wong et al., 1997).

Alternatively, harvested cells may be lysed by chemical methods. Chemical lysis methods tend to be harsh and can solubilize not only cell wall/membrane, but the IBs as well. A method has been developed to extract IB recombinant protein directly from cells (Falconer et al., 1998), but the purity of protein recovered is considerably lower than that isolated by mechanical methods. A lysis technique allowing release of unsolublised IB has been established (Falconer et al., 1999), producing recoveries and purity comparable to conventional mechanical disruption procedures, and can be applied *in-fermenter* (Lee et al., 2004). Pierce Chemical company (Rockford, Illinois) also have commercially produced a non-solubilising bacterial extraction kit for IBs (B-PER) (Ruiyin and Krishna, 2001).

1.2.4.3 IB Isolation

Extracted IBs not only have to be separated from contaminants in the soluble phase but also from cell debris present in the solid phase. Repeated rounds of homogenisation can reduce cell debris size considerably, without affecting the size of IB, but it is likely that the IB and cell debris size distributions overlap to a certain extent. Separation is therefore best achieved using centrifugation by virtue of the greater density of the IB (Middelberg, 2002). At industrial scale this separation is commonly achieved in disc-stack centrifuges, where adjustment of flow rate, can favour the sedimentation of inclusion bodies over cell debris (Jin et al., 1994). Recovery of IBs generally increases with decreased flow rate, but results in poor removal of cell debris, whilst high flow rates result in lower recovery but higher purity of IBs. A compromise must therefore be sought (Hoare and Dunnill, 1989). This separation will be by no means absolute, and cell debris is likely to be associated with the IB. Microfiltration has been used with some

success to isolate IB from chemically lysed cells (Lee et al., 2004), but may be unsuitable for mechanically isolated IB because selectivity is based only on size.

1.2.4.4 Washing of IB

Isolated IBs are typically contaminated with cell debris. Contaminating cell debris has been shown to: increase the degree of proteolysis of IB (presumably through the presence of proteolytic enzymes) (Wong et al., 1996), affect the yield from refolding (Georgiou and Valax, 1999; Maachupalli-Reddy et al., 1997) and may present an extra burden on downstream processing steps (Thatcher et al., 1996). Cell debris consists of the insoluble matter in the cells i.e. cell walls, membranes etc. IB preparations are typically contaminated with peptidoglycans, membrane proteins, lipid and nucleic acids (Thatcher, 1990). To remove these non-specific contaminants IB washing steps may be employed. These washing stages may be associated with centrifugation or microfiltration. In centrifugal washing IBs are resuspended in wash buffer, mixed for a short period, spun down and the supernatant removed. Resuspension marks the next step. A comparison of centrifugal and microfiltration washing procedures, showed that flux and protein purity were achieved with larger pore sizes in MF, but centrifugation produced IB of higher purity overall (Batas et al., 1999).

Typical washing additives include Sucrose (1 M) (Sugimoto et al., 1991; Sugimoto and Yokoo, 1991), Triton X (0.1-4% v/v) (Batas et al., 1999; Buswell et al., 2002; Sugimoto et al., 1991; Sugimoto and Yokoo, 1991), urea (2-5 M) (Batas et al., 1999; Fromage et al., 1991; Smith et al., 1990), deoxycholate (Langley et al., 1987), and in some cases enzymes such as lysozyme (Buswell et al., 2002).

1.2.4.5 Solubilization

After the inclusion body has been purified it is essential that it is solubilised and the protein within it unfolded. This process disassociates the aggregated polypeptides and leaves polypeptide in an unfolded state ready for refolding.

Solubilization is most commonly achieved by resuspending the IBs in an appropriate chaotrope, in the presence of reducing agents.

Concentrated solutions of urea (8M) or guanidine hydrochloride (GdHCl) (6-8M) are typically used (Fischer et al., 1993). Other methods of solubilzation/denaturation include extremes of pH, detergents such as SDS and CTAB, lithium chloride, potassium thiocyante, calcium chloride, guanidine thiocyanate, sodium perchlorate, high temperatures, and high hydrostatic pressures (Rudolph et al., 1997). Guanidine hydrochloride and urea are usually favoured because of their high denaturing capacity, with the exception of guanidine thiocyanate which is stronger denaturant, but is less commonly used. Combinations of denaturant can be used to enhance the denaturing effect. Care must be taken when denaturing proteins using extremes of temperature and pH as this can cause chemical alterations, such as hydrolysis of peptide bonds, oxidation and deamidation (Rudolph et al., 1997). Care should also be taken when using some denaturants, in particular urea. Urea decomposes into isocyanate, which can carbamylate the free amino groups of the polypeptide. This effect is particularly evident when it is incubated at alkaline pH for long periods (Rudolph et al., 1997) Choice of denaturant can be critical to the recovery of active protein in subsequent refolding steps. Solubilization in GdHCl increased the recovery of human interleukin 4, whereas solubilization in urea recovered no active protein (Van Kimmenade et al., 1988).

Most inclusion body solubilizations require the presence of reducing agents to break down non-native intramolecular and intermolecular disulfide-bonds (Fischer et al., 1993). Typically 2-mercaptoethanol or dithiothreitol (DTT) are used as reducing agents, but monothiols such as cysteine may also be used (Buswell et al., 2002). The dithiol DTT can be used at lower concentration than monothiol because of its higher redox potential (Fischer et al., 1993). Solubilization can be achieved without reducing agents. In such situations the protein in IB may contain no intermolecular or intramolecular disulfide bonds. It was observed for bovine growth hormone that the protein was in a completely reduced state (Langley et al., 1987), and therefore denaturation could be achieved in the absence of reducing agent. Despite denaturation being possible in the absence of reducing agents it is advisable to include them to ensure full denaturation/solubilization occurs.

Denaturation and reduction conditions are usually chosen empirically. The denaturant, reducing agent, temperature and period of denaturation must be established, to enable full denaturation to occur.

Denatured solubilised protein is then ready for refolding. To understand the factors affecting the refolding of proteins, it is essential to gain at least a basic knowledge of protein structure and the molecular interactions involved in folding.

1.3 The Process of Folding

1.3.1 The Four Levels of Protein Structure

The diversity of size, shape and the inherent chemical reactivity of the main 20 amino acids, determines the way a polypeptide folds (Gardiner, 1988). There are

four levels of protein structure, whose increasing structural complexity is brought about by folding of protein structure, via molecular interactions as follows:

- Primary- this is the polypeptides linear sequence of amino acids. It is formed through the formation of peptide bonds between amino acid residues upon translation.
- Secondary – this is the structure formed through interactions between residues close together in the amino acid sequence. Examples of secondary structure include the α-helix(Pauling et al., 1951) and β-pleated sheets (Pauling and Corey, 1951).

  In the cylindrical alpha helix the contributing amino acids arrange themselves in a helical fashion (Figure 1.3). The helix is formed by hydrogen bonds, between the carbonyl-oxygen of each peptide bond and the hydrogen atom of the amino group, four amino acids along the chain (Stryer, 1995). This produces a structure that has 3.6 amino acids per a turn of the helix (Stryer, 1995). In this structure the amino-acid side chains are on the outside of helix. Amino acids are found in differing proportions in alpha helices. Of particular interest is proline. The absence of a hydrogen atom on the nitrogen of this amino acid prevents it forming the hydrogen bonds essential for the formation of the alpha helix. As a consequence proline is often found at the end of an alpha helix where it alters the direction of the polypeptide. The protein myoglobin (18 kD), the monomer of the tetramer haemoglobin is composed primarily of alpha helices, of which it contains eight (Stryer, 1995).

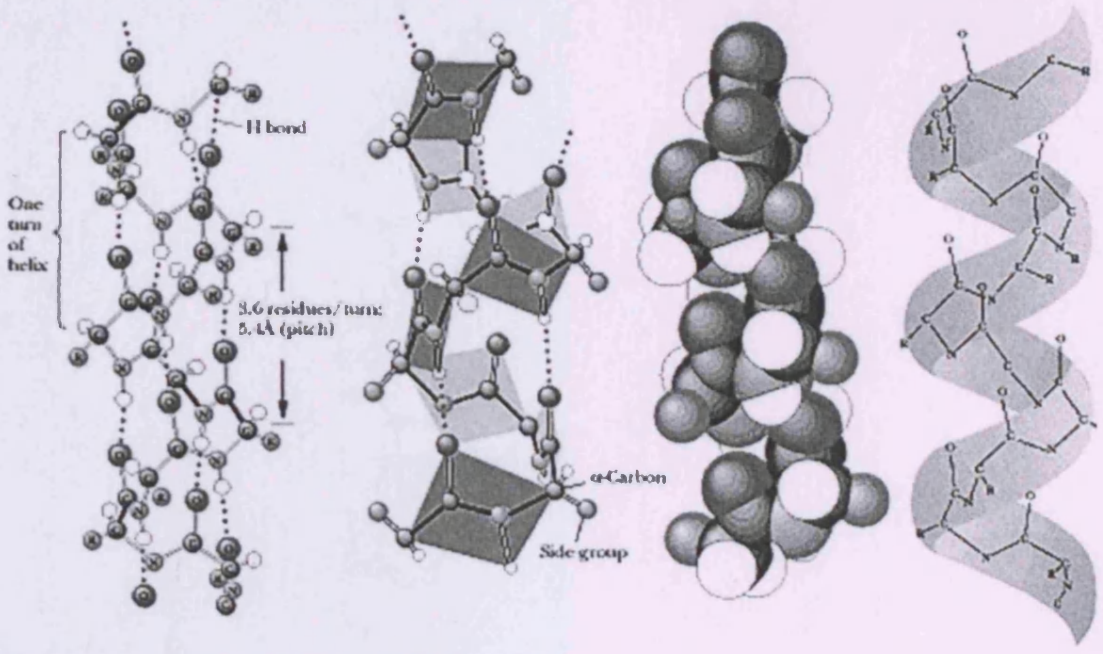


Figure 1.3: Various models of the alpha helix, showing the hydrogen bonds required and structure produced through their action. Figure from (Campbell, 1999).

Like the alpha helix the beta pleated sheet is stabilised by hydrogen bonds. These hydrogen bonds are formed between different polypeptide chains or between different sections of the same polypeptide (Hames et al., 1997). The planar nature of the peptide bond provides this structure with its pleated character, where the amino acid side chains protrude above and below the sheets in a consecutive fashion (Hames et al., 1997). The pleated nature of this element of secondary structure can be seen in Figure 1.4. The adjacent polypeptide chains or sheets are described as parallel or anti-parallel, according to their orientation to the adjacent polypeptide chain.

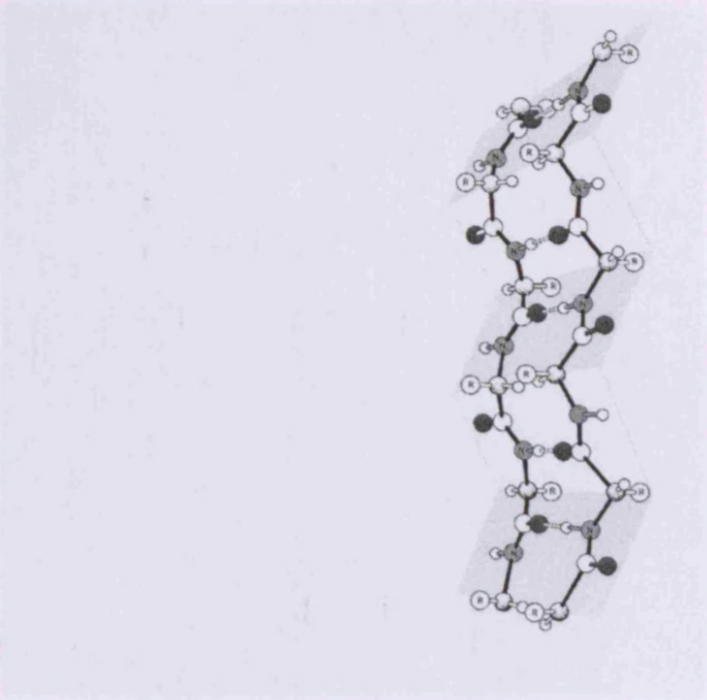


Figure 1.4: Ball and chain model of the beta-pleated sheet. Figure from(Campbell, 1999)

- Tertiary structure- formed through interactions between regions far apart in the amino acid sequence of the protein, these interactions may be facilitated through the coming together of domains (ordered local structures). If the protein is monomeric, this conformation is the native structure of the protein

- Quaternary structure- in polymeric proteins is the association of monomers to form an oligiomeric structure. The structure of the monomer in its oligiomeric state is similar if not identical to the monomer in its monomeric state (Gardiner, 1988) (Figure 1.5).

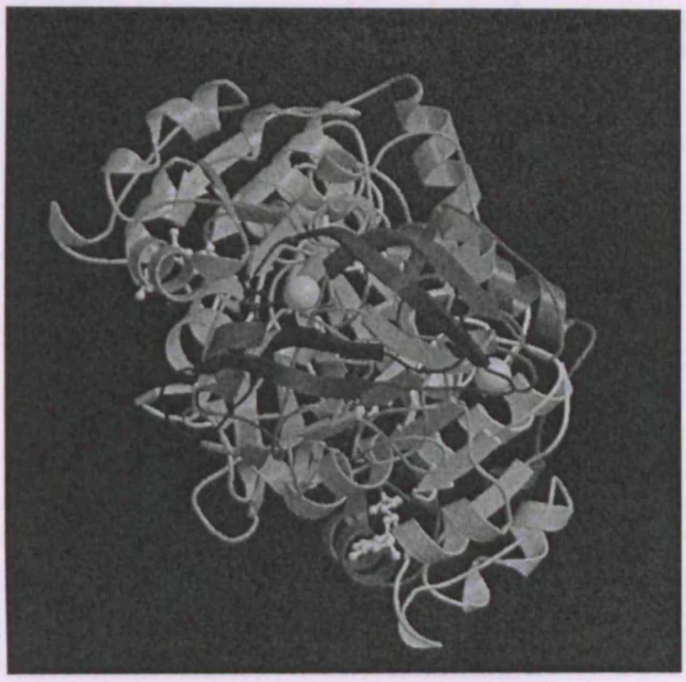


Figure 1.5: Quarternary structure of liver alcohol dehydrogenase, with bound NAD and zinc from (Al-Karadaghi et al., 1994).

Molecular interactions are key to allowing the formation of structures described above.

1.3.2 The Molecular Interactions Permitting Folding

The rigidity and planar nature of the peptide bond limits the freedom in folding. The peptide does, however, possess a limited rotational freedom about the carbonyl α-carbon and nitrogen α-carbon bond between each successive peptide bond (Figure 1.6) (Gardiner, 1988).

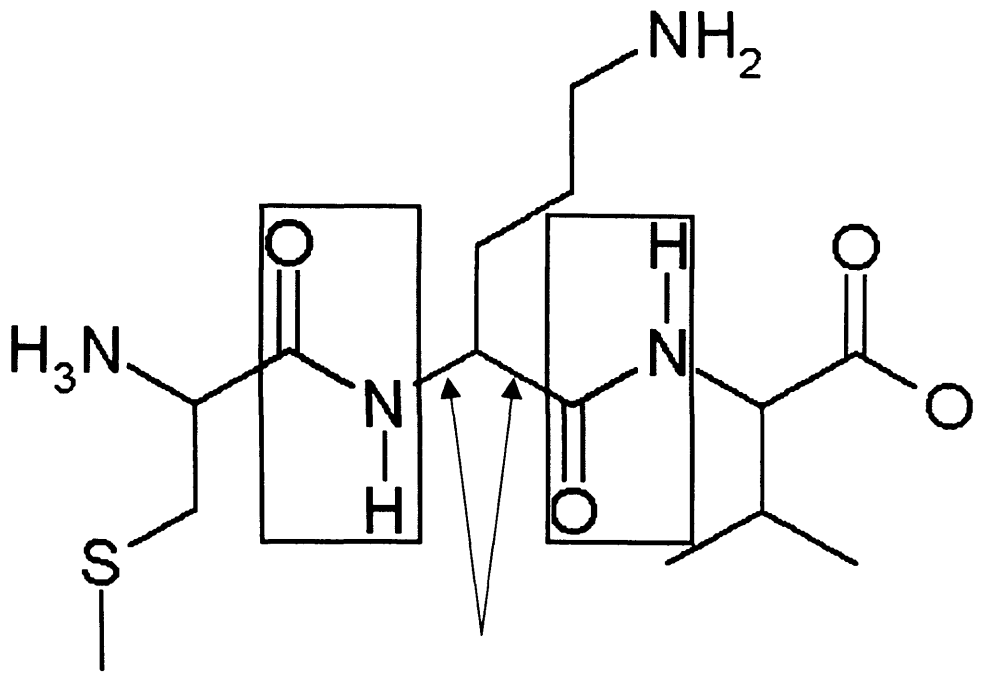


Figure 1.6: Planar nature of peptide bonds. The atoms in the box (peptide bond) form a rigid planar unit, whilst the two bonds indicated by arrow show a high degree of rotational freedom. (Adapted from (Alberts et al., 1994)).

1.3.2.1 Hydrogen Bonds

Hydrogen bonds are important structural interactions involved in the folding of protein. These interactions readily form between the CO and NH groups in the polypeptide backbone; this interaction permits the formation of α-helices and β-pleated sheets. Formation of alpha helices and beta pleated sheets occur with growth times of 10^{-8} s^{-1} and 10^{-2} s^{-1}, respectively (Narayana and Argos, 1984). It is therefore thought that such interactions occur early in the folding process. Hydrogen bonds provide an important contribution to protein stability with bonds strengths of between –8.4 and 33.5 $KJmol^{-1}$ (Ghelis and Yon, 1982).

1.3.2.2 Electrostatic Interactions

Electrostatic interactions are also critical in the folding of proteins. At any given pH, many of the amino acid side chains are charged (even at the isoelectric point of the protein (Gardiner, 1988). When these charged groups are located in the interior of the protein, opposite charges will readily pair to minimise the overall charge on the protein (Chothia, 1976). Charges that are unpaired are likely to be located at the surface where they may interact with the solvent. Properties of the solvent, such as ionic strength for aqueous media and dielectric constants for non-polar solvents, determine the strength of such interactions (Gardiner, 1988). Electrostatic interactions, have bond strengths –4.2 to 12.5 KJ Mol^{-1} (Kim and Baldwin, 1982), those located in the interior of the protein are more important to protein stability, than those existing at the surface of the protein (Gardiner, 1988).

1.3.2.3 Hydrophobic Interactions

The function of hydrophobic groups in the stabilisation of protein has been extensively studied (Burley S K and Petsko, 1985; Kuharski and Rossky, 1984; Némethy and Scheraga, 1962a; Némethy and Scheraga, 1962b; Rose et al., 1985). A number of the amino acid side chains are hydrophobic. Hydrophobic residues proportionality make up 0.35- 0.8 of residues in globular proteins, dependent on the classification of side chains as hydrophilic or hydrophobic (Kauzmann, 1959; Tanford, 1968). It is energetically unfavourable for such moieties to associate with water; hence, they locate themselves away from water in the interior of the protein. Such hydrophobic moieties interact to stabilise the structure of the protein. The typical strength of these hydrophobic interactions are –4.2 to –8.4 KJ mol^{-1} (Burley S K and Petsko, 1985). Unfolding requires destabilisation of

these interactions, and the only way this may be achieved is by making the transfer of hydrophobic group into an aqueous environment energetically favourable (Gardiner, 1988).

1.3.2.4 Disulfide Bonds

Of particular importance in the stabilisation of protein structure is the formation of disulfide bonds. These covalent bonds result from bonding between cysteine residues, upon oxidation of their thiols. The bond energy of a disulfide bond is 250 kJ mol-1 (Gilmore, 1977), and provides considerable stabilisation. Studies suggest that disulfide bonds are more likely to form between cysteine residues close together in the amino acid sequence (Ghelis and Yon, 1982), than between those located further apart which may permit domain association. It appears that there is no clear relationship between protein size and the number of disulfides present (Ghelis and Yon, 1982). As the number of cysteines in structure increases, the number of possible pairings increases considerably, however only a particular set of pairings will form the native structure.

1.3.3 How Does Folding Occur?

The previous section has provided an introduction to the molecular interactions that allow folding. This section tries to understand how proteins fold.

1.3.3.1 Initial Studies

As early as the 1920s protein denaturation and recovery to the native state were studied (Wu, 1931). Real clues to how protein folded were established in refolding studies in the 1960s on ribonuclease A (Anfinsen and Haber, 1961;

White, 1961). Ribonuclease A is a 124 amino acid protein that catalyses the hydrolysis of phosphodiester bonds in the degradation of RNA. In these experiments ribonuclease was reduced and denatured in 8M Urea in the presence of β-mercaptoethanol. The denatured reduced protein formed a randomly coiled chain devoid of any activity. Upon dialysis to remove the β-mercaptoethanol and urea, the ribonuclease slowly regained activity. These experiments provided the evidence for the thermodynamic hypothesis, which suggests that the native structure of a protein in its normal physical milieu is the state in which the Gibbs free energy is the lowest. This native structure is determined by all the inter-atomic interactions and therefore by the amino acid sequence (Anfinsen, 1973). In essence all the information required to fold a protein is encoded with in its amino acid sequence (Stryer, 1995).

How does folding occur? Are all the possible conformations searched before the native is found or is the process more directed? Cyrus Levinthal addressed this problem (Levinthal, 1968). He reasoned that for a protein of just 100 amino acids there are 3^{100} ($5x10^{47}$) different possible structures. Assuming that it takes $1x10^{-13}$ s to swap between conformations it will take $5x10^{34}$s or $1.6x10^{27}$years to search for the correct structure. This suggests that folding is not achieved through a random search. The difference between the time calculated for random selection and the actual refolding is called Levinthal's paradox (Stryer, 1995).

It is therefore believed that a more cumulative selection process exists, such that partially correct intermediates are retained. It is important to note that there is only a small difference between the free energies of the unfolded and folded

states. For an average 100 amino acid protein this difference is just 10 kcal/mol (Stryer, 1995). This equates to just 0.1 kcal stabilisation per a residue.

1.3.3.2 Refolding Intermediates

If a directed process of folding occurs it is likely that transient intermediates occur. The transiency and co-operativity of such intermediate structures has made them particularly difficult to isolate. Theoretical considerations suggest that folded flexible intermediates exist in the refolding process (Ptitsyn and Rashin, 1973), with the name "molten globule" being coined (Ptitsyn and Rashin, 1973). Such species are compact with both a high content of native secondary structure and transient tertiary structure. Such molten globule species have been observed for α-lactalbumin (Kuwajima, 1989) and carbonic anhydrase (Jagannadham and Balasubramanian, 1985) amongst other proteins. It has been suggested that the molten globule is a general intermediate in the folding of proteins (Ptitsyn et al., 1990). Evidence from studies on α-lactalbumin, suggest that there is a heterologous mixture of molten globules (Kuwajima, 1996).

Intermediates do not necessarily contain native secondary structure. The refolding of lysozyme (Radford et al., 1992) requires transient steps to change secondary structure in the intermediate to derive the native form.

It is thought that there are species intermediate to unfolded and molten globule structures (Uversky and Ptitsyn, 1994; Uversky and Ptitsyn, 1996). These pre-molten globules are less compact than the molten globule, having hydrophobic regions accessible to solvent and with a degree of secondary structure but less than the molten globule. This suggests that folding occurs from secondary structure interactions.

An alternative view of refolding suggests that hydrophobic collapse is the earliest stage perhaps even preceding the formation of secondary structure. Experimental evidence appears to support this (Amir et al., 1992; Garcia et al., 1995; Radford et al., 1992). Evidence from experiments allowing measurement of folding events to the nanosecond time scale (Ballew et al., 1996) have suggested that the initial stages of protein refolding are hydrophobic collapse accompanied by secondary structure formation. It is thought that molten globules form after these fast events. These are then followed by further intermediate steps. The rate determining steps are thought to involve precise ordering of the secondary structure elements, and correct packing of the hydrophobic core (Matouschek et al., 1990). In some cases this may involve cis-trans isomerisation of proline residues (Levitt, 1981) and disulfide shuffling (Wetlaufer and Ristow, 1973). In multi-domain and multi-subunit proteins, domain and subunit association also limit the rate of refolding.

1.3.3.3 Towards a Unified Model

From the wealth of experimental data it was generally accepted by scientists in the field until the mid 90s that folding occurred in a sequential manner, as detailed in Figure 1.7:

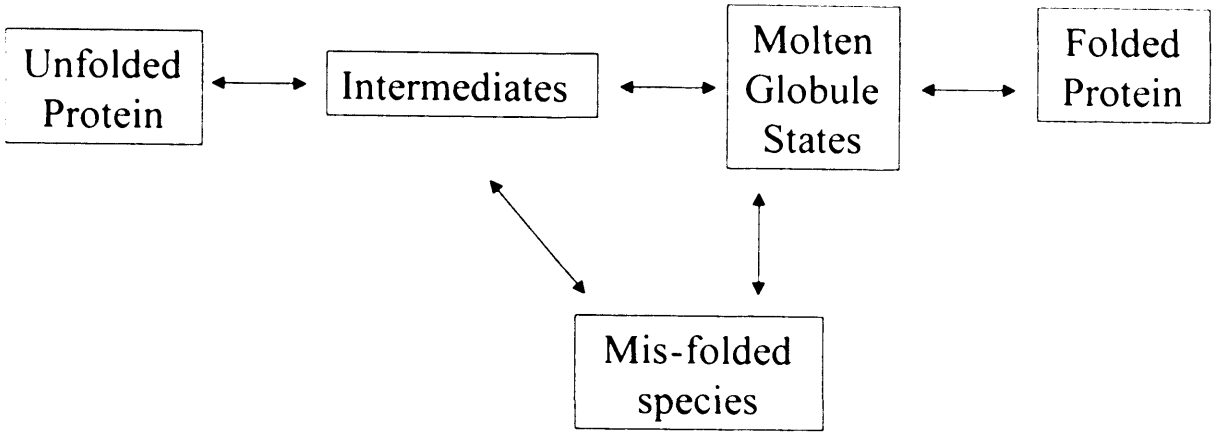


Figure 1.7: Summary of sequential folding models (adapted from (Yon, 2002)).

Recent theoretical work, has generated a unified model, which looks at the effective energy surface of polypeptide chain (Wolynes et al., 1995). This method describes folding in terms of an energy landscape and a folding funnel. The model portrays both thermodynamically and kinetically the behaviour of the conversion of a collection of unfolded molecules to a final predominantly native state (Yon, 2002). Such a collection of partially folded intermediates are derived through a multitude of routes, which gradually organise over time (Brygelson et al., 1995; Onuchic et al., 1995; Onuchic et al., 2000). This description makes Levinthal's paradox less of a problem, as the protein has a multitude of routes from which to form its native structure (Yon, 2002). Experimental evidence appears to support this theory where parallel alternative pathways have been observed for both lysozyme and cytochrome c (Dobson et al., 1998).

Progressing down the refolding funnel to the native conformation, the number of conformations decreases, as does the chain entropy. The steeper the slope the faster the reaction (Yon, 2002). An example of such a refolding funnel is shown schematically in Figure 1.8. If the funnel were completely smooth this would represent a very simple two state folding event where native structure is attained

quickly. This, however, is rarely the case (Yon, 2002). Very typically refolding funnels are rough, the roughness corresponds to kinetic traps arising from energy barriers. Such a refolding process will be slow and multi-state. If energy barriers are high enough molecules may become trapped in an intermediate state and may aggregate. Typically when considering kinetic processes, transition states are typically thought of as a single structure. In the case of the "funnel" description of protein refolding this transition state may be thought of as a distribution of structures (Yon, 2002).

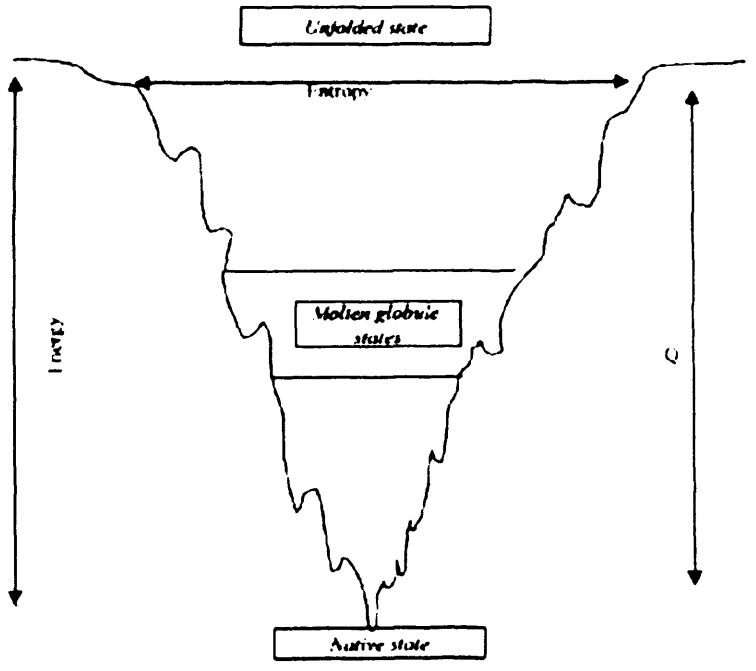


Figure 1.8: A 2-dimensional representation of a folding funnel, with corresponding refolding states indicated. Entropy is indicated as the width of the funnel whilst energy is the height. Q represents the fraction of native contacts for each state. Taken from (Yon, 2002)

1.4 Protein Refolding

The previous section provided a guide protein structures and folding. This section provides an insight into refolding; the process by which active protein is generated from purified, solubilised IBs, where folding is achieved *in-vitro*.

1.4.1 The Process of Refolding

After protein has been unfolded and reduced, a process called refolding derives active protein. Refolding has two requirements: First the removal of chaotrope, and for disulfide-containing proteins, a redox environment which permits efficient disulfide exchange. Unlike folding *in-vivo*, cellular components such as chaperones are not present to prevent misfolding and aggregation. Conditions for refolding are hence chosen to minimise unproductive pathways, but maximise the amount of protein refolded.

1.4.1.1 Chaotrope Removal

The conformation, flexibility and solubility of a protein is a function of chaotrope concentration (Figure 1.9) (Tsumoto et al., 2003). Refolding is achieved by removing chaotrope from the protein, which allows collapse of structure, permitting refolding, but it may also induce aggregation. There are various methods by which this state may be achieved. These are outlined in section 1.4.3.

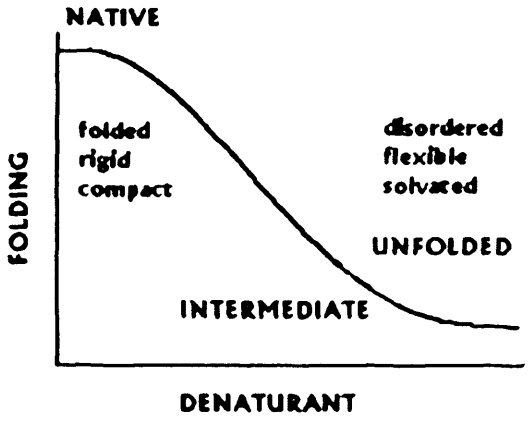


Figure 1.9: Effect of denaturant concentration on form of protein (from (Tsumoto et al., 2003). This figure slightly simplifies the situation as protein solubility can also decrease with denaturant concentration, and aggregation rather than refolding can result.

1.4.1.2 Disulfide Exchange

If the native protein contains disulfide bonds it is critical that the refolding buffer has redox reagents, which permit disulfide exchange. This is most commonly achieved using 'oxido-shuffling' systems (Rudolph et al., 1997). Such buffers use mixtures of thiol (RSH) and disulfide (RSSR). Thiol species reduce disulfide bonds whilst disulfides reform disulfide bonds. Such a system permits formation, breakage and reformation of disulfide bonds, permitting correct disulfide pairings to be found (Figure 1.10).

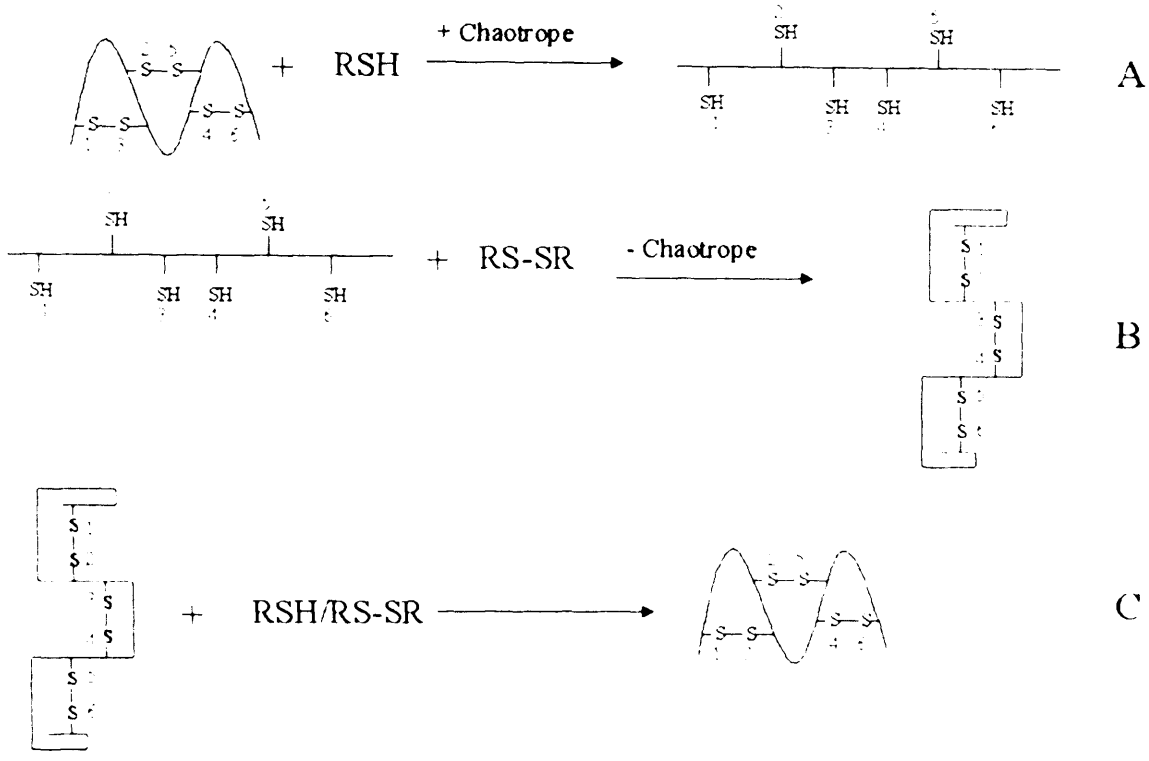


Figure 1.10: A description of how oxido-shuffling systems operate. (A) Protein is denatured and reduced by use of chaotrope and reducing agent (RSH). (B) Removal of chaotrope allows molecular collapse, with oxidising redox species (RS-SR) allowing oxidation of disulfides. (C) Oxidation may produce non-native disulfide pairings a mixture of oxidising and reducing species can permit shuffling until the correct pairings are found.

Typical redox pairs (reduced:oxidised) include: reduced glutathione:oxidised glutahione, cysteine:cystine, cysteamine:cystamine, DTT:oxidised DTT, and mercaptoethanol:2 –hydroxyethyldisulfide (Rudolph et al., 1997).

Alternatively, reduced denatured protein can be renatured in buffer by air oxidation, where air is used to oxidise reduced disulfides. In such cases the refold buffers used contain low concentrations of metal ions such as Cu^{2+} (typically micromolar levels) (Rudolph et al., 1997) to catalyse the oxidation process. Such methods are better suited to refolding using dilution methods, where air may be incorporated.

A common problem with refolding is the instability or insolubility of reduced protein (Rudolph et al., 1997). This issue may be circumvented by the reversible modification of protein thiols, with the introduction of charged species increasing solubility (Rudolph et al., 1997). Modification can be achieved via sulfonation by treating unfolded species with 2-nitro-5-(thiosulfo)benzoate (NTSB), which cleaves disulfide bonds producing sulfonated thiols. Refolding is then achieved using a oxido-shuffling system. Alternatively, denatured reduced protein can be treated with high concentrations of oxidising agent to form mixed disulfides, which can be refolded in buffer containing reduced redox reagent.

1.4.1.3 Refolding Buffers

Removal of chaotrope is usually achieved in refolding buffers. There is no universal buffer permitting the successful refolding of all proteins from their denatured states, a reagent in one refolding buffer may permit successful refolding of one protein, but may not permit it of another or worse it may induce aggregation. Determination of the optimum conditions to be used must be achieved empirically. A variety of agents can be added to improve refolding yields, these are discussed in section 1.4.10.

1.4.2 The Competition Between Refolding and Aggregation

Refolding is not a single reaction. It competes with misfolding and aggregation (Figure 1.11) (Tsumoto et al., 2003). The refolding and aggregation reactions compete for a folding intermediate. Refolding is a monomolecular reaction, involving intramolecular reactions and can be described by first-order kinetics,

whilst aggregation is multi-molecular and is described by higher order kinetics (De-Bernardez Clark et al., 1998). Competition exists between intra-chain interactions permitting refolding, and inter-chain reactions causing aggregation. Given the competition and order of the reactions the formation of native protein is favoured at low protein concentrations (Buswell et al., 2002).

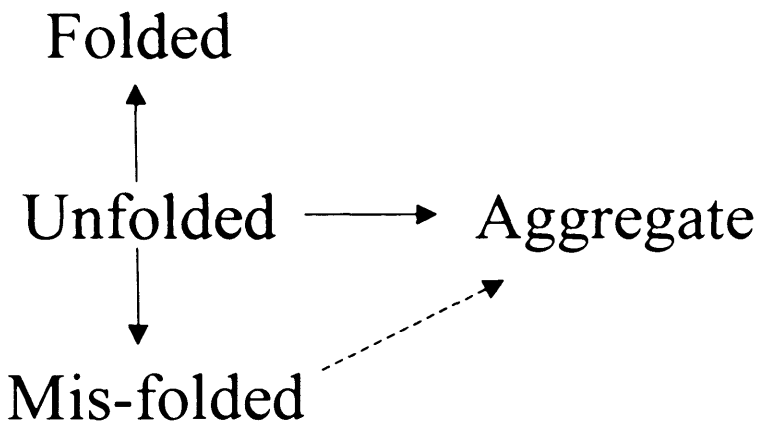


Figure 1.11: A simplified summary of refolding and its competing reactions. Mis-folded protein can accumulate in aggregates.

Aggregation is the major cause of losses in yield from refolding reactions. Aggregation is a process whereby thermolabile folding intermediates associate. It may be considered as a polymerisation reaction, but it is not a crystallization process (Buswell et al., 2002). The extent of aggregation is primarily determined by the protein concentration, where the likelihood of aggregation increases with protein concentration. The degree of aggregation appears to be determined by denaturant and protein concentration for the refolding of both carbonic anhydrase B (Cleland and Wang, 1990b) and lactate dehydrogenase (LDH) (Rudolph et al., 1979; Zettlmeissl et al., 1979). It has been observed that for several proteins a critical denaturant concentration exists (Mitraki and King, 1989). At this critical

concentration the rate of aggregation of the intermediate in the refolding reaction may exceed that of its refolding, resulting in low recovery of active protein (Cleland and Wang, 1993). It is therefore thought that the degree of aggregation is primarily affected by the concentration of protein and denaturant present.

Non-covalent interactions are the most likely driving force for aggregation (Cleland and Wang, 1993). It is theorized that hydrophobic interactions drive aggregation. The kinetically endothermic nature of hydrophobic interactions (Mitraki and King, 1989), concomitant loss of entropy upon aggregation (Ghelis and Yon, 1982) and the exposure of hydrophobic regions by chaotropic agents (Mitraki and King, 1989), suggest aggregation is caused by hydrophobic interactions (Cleland and Wang, 1993).

Furthermore, studies on egg white lysozyme suggest that incorrect disulfide bonding may not be the major cause of aggregation (Goldberg et al., 1991). In this study, cysteines were blocked with carboxymethyl groups to prevent disulfide bonding. Refolding of this protein still resulted in aggregation, suggesting disulfide bonds are not critical to aggregation.

The specificity of aggregation has been subject to several studies. Evidence from Goldberg et al suggested that aggregation was non-specific (Goldberg et al., 1991). However, studies of refolding P22 tailspike and coat protein, showed that these proteins did not aggregate together, but preferred to self associate, suggesting specificity (Speed et al., 1996). Further evidence for the non-specificity of aggregation comes from mixed protein studies (Maachupalli-Reddy et al., 1997). It was found that BSA and β-galactosidase, which aggregate when folded in isolation, appeared to effect lysozyme refolding yields in co-folding studies.

Studies on lysozyme refolding suggest that aggregate concentration does not increase significantly after the first minute of the refold, where hydrophobic interactions, and not disulfide bonds, caused aggregation (De-Bernardez Clark et al., 1998). It was found that disulfide bonds were involved in increasing aggregate size but not concentration. It is therefore suggested that small soluble aggregates may be formed through hydrophobic interactions between the same species but that large heterogeneous aggregates may be formed through non-specific disulfide bonding (De-Bernardez Clark, 1998).

It is the aim of the refolding methods to limit the extent of aggregation. Typically refolding methods will use low concentrations of protein upon refolding to limit aggregation, unless spatial separation can be achieved by an alternative method. Industrially, low concentrations are undesirable as this requires processing of large volumes. A compromise therefore exists between the concentrations used and the need to keep equipment size minimal.

1.4.3 An Overview of Protein Refolding Methods

Various methods exist for refolding proteins, broadly categorised into 3 groups:

- Dilution refolding methods
- Buffer exchange based methods
- Chromatographic methods.

In the following sections the advantages and disadvantages of each technique will be described, their suitability for application in industrial processes contrasted, and the ability to perform multiple experiments with each technique compared. In addition a description of new novel methods is also included to provide an insight into new techniques.

1.4.4 Dilution Refolding Methods

A commonly utilised method of refolding denatured protein from inclusion bodies is refolding by dilution (Buchner and Rudolph, 1991; Buswell et al., 2002; Davis et al., 1987; Halenbeck et al., 1989; Lee et al., 2002; Van Kimmenade et al., 1988). Denatured protein in high concentrations of chaotrope are diluted into large volumes of buffer commonly containing redox reagents to effect disulfide exchange (Tsumoto et al., 2003). Using dilution refolding, denaturant and protein concentration quickly go from high to low levels, inducing a rapid collapse in the structure of the protein to a rigid form, which may or may not allow conversion to the native state (Figure 1.12A) (Tsumoto et al., 2003). To alleviate this effect it is integral for many proteins that a low concentration of denaturant is included in the refold buffer to permit some molecular flexibility (Tsumoto et al., 2003).

Refolding by dilution is most successful at low protein concentrations, as the likelihood of aggregation is minimised. Aggregation results in significant losses in the amount of native protein. This is a major limit on the feasibility of recombinant protein manufacture by the IB route (Buswell et al., 2002). Refolding at low concentrations is, however, impractical at industrial scale; alternatives permitting more efficient refolding of larger amounts of protein must therefore be sought. Such a method is the patented pulsed renaturation method (Rudolph and Fischer, 1990). In this method the denatured protein is added in pulses at a critical protein concentration. After each pulse the solution is left to allow refolding, and then the next pulse is added. This process permits high amount of protein to be refolded, as it keeps the effective denatured/intermediate protein concentration within the refolding reaction to a minimum, reducing the potential for aggregation.

Other methods exist for dilution refolding, including reverse dilution (Figure 1.12B) where refolding buffer is added gradually to the denatured protein allowing a gradual decrease in denaturant concentration. This method, however, suffers from the presence of high concentrations of protein at intermediate denaturant concentrations, which can result in aggregation (Tsumoto et al., 2003). It is not commonly used at scale.

Mixing dilution (Figure 1.12C) is not commonly used. Here denatured protein and refold buffer are mixed in constant proportions. This method is similar to dilution refolding as it induces rapid protein structure collapse (Tsumoto et al., 2003).

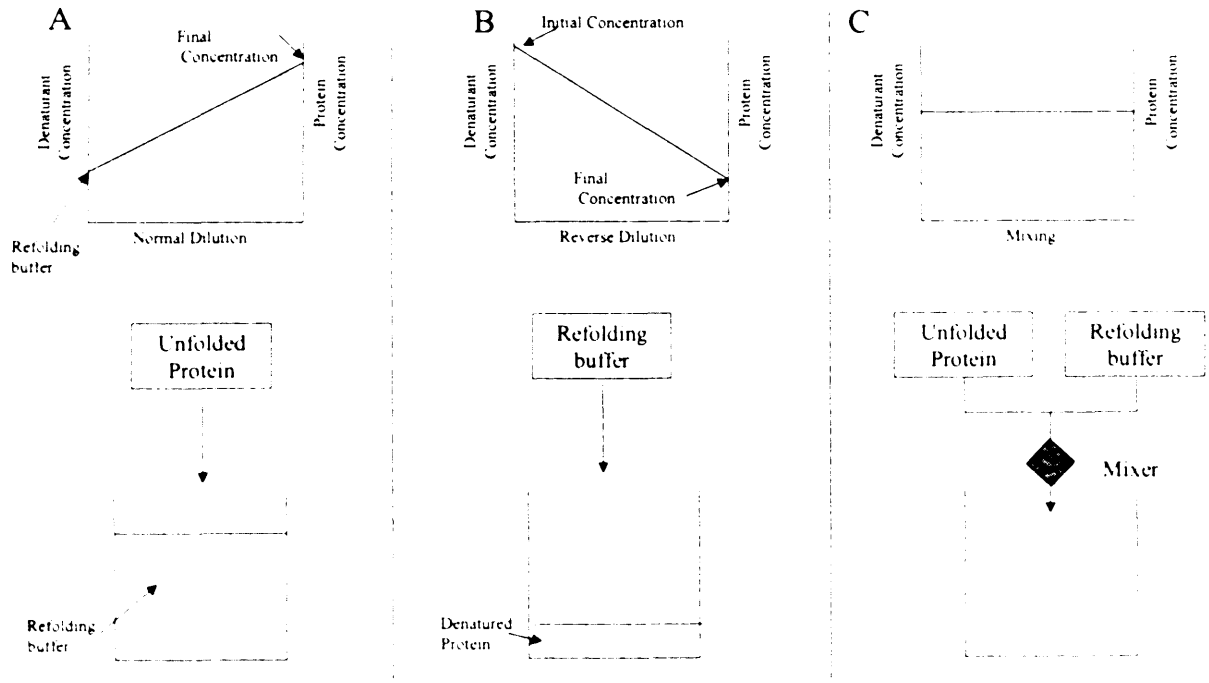


Figure 1.12: A summary of dilution refolding methods. Top line shows the denaturant/protein concentration over the course of the refold. Bottom line is a schematic of each of the dilution refolding methods (A) Dilution refolding, (B) Reverse dilution, (C) Mixing dilution. Adapted from (Tsumoto et al., 2003).

1.4.5 Buffer Exchange Methods

Another means of refolding is buffer exchange. This is most commonly achieved by diafiltration) (Varnein et al., 1998) or by dialysis (West et al., 1998). In such methods the denaturing and reducing buffer is exchanged for refolding buffer,

either by diffusion against refold buffer or by transmembrane pressure, and addition of refold buffer in the case of diafiltration. The process in some respects is akin to the reverse dilution process where the concentration of denaturant is gradually reduced. The difference here is that protein concentration in most instances stays constant, as protein should be unable to traverse the membrane. It is likely that the protein will be subject to long periods at intermediate denaturant concentrations and this may well prompt aggregation especially if the rate of refolding is slow, as denatured/ intermediate forms may not be soluble at lower concentrations of denaturant (Tsumoto et al., 2003). A critical parameter in such operations is protein concentration, which must be kept low enough to prevent aggregation.

To improve refolding yields from dialysis/diafiltration methods, the process can be achieved by a step dialysis/diafiltration. Under this protocol the protein is equilibrated at a series of decreasing denaturant concentrations. This method has been successful in the refolding of some antibodies (Tsumoto et al., 1998). Success of such a method is dependent on the rate of aggregation/misfolding being slower than refolding and this may require refolding at quite low concentrations (Tsumoto et al., 2003). It may be of use for the refolding of multi-domain proteins, where the stability and refolding of domains is different (Tsumoto et al., 2003).

A particular problem encountered with refolding using membranes for buffer exchange, is the binding of protein to the membrane (De-Bernardez Clark, 2001). To reduce these affects it is preferable to use membranes made of highly hydrophilic material such as cellulose acetate, as opposed to more hydrophobic materials such as polyethersulfone, where large amounts of the denatured protein

have been shown to bind (West et al., 1998). Additionally, significant losses occurred through transmission of the denatured protein across the membrane, which is more rapid than for native globular forms (West et al., 1998).

1.4.6 Chromatographic Methods

Refolding by chromatography is an attractive means of refolding as it may be automated utilising existing rigs (Middelberg, 2002). There are essentially three different types of refolding chromatography available: Solvent exchange by size exclusion chromatography; adsorption refolding; and use of immobilised catalysts (Li et al., 2004).

1.4.6.1 Size Exclusion Chromatography

Protein may be refolded by size exclusion chromatography (Batas et al., 1999; Werner et al., 1994). The nature of size exclusion chromatography means the volume of the matrix available to a protein is determined by its size. This permits the separation of larger aggregate molecules from correctly folded protein (Li et al., 2004). Denatured protein can be loaded onto a column equilibrated with either refold buffer or a gradient of denaturant. Use of a gradient of denaturant can improve yields considerably (Batas et al., 1999). High yields can be achieved using this method, but the protein undergoes considerable dilution (up to 50-fold have been reported) (Middelberg, 2002). Little advantage is gained in use of SEC over dilution methods at comparable final protein concentrations though its ability to separate molecules of different size means that the refolded protein derived may be relatively free of contaminating proteins and misfolded/ aggregated protein forms (Middelberg, 2002).

1.4.6.2 Adsorption Chromatography

Adsorption chromatography uses adsorptive resin such as ion exchange (IEX), hydrophobic interaction (HIC) and immobilised metal affinity (IMAC) to adsorb the denatured protein. By binding to the matrix denatured protein molecules are spatially separated from each other preventing aggregation. Denaturant is then diluted to allow refolding of proteins (Middelberg, 2002). The applicability of adsorptive chromatographic techniques is incredibly protein specific, with some protein-matrix interactions preventing refolding (Middelberg, 2002). Studies showed that horse cytochrome could be absorbed from urea onto an ion exchange column and subsequently refolded (Creighton, 1986). Absorbance can be specific or non-specific; though specificity is preferred. Specific binding to matrices may be achieved through specific domains. Such binding will ideally leave the majority of the protein free to refold. Specific interactions may be achieved by protein modification. In such a study a six arginine residue tail was added to alpha-glycosidase to effect specific immobilisation on a polyanion matrix (Stempfer et al., 1996). From this study it was shown that refolding conditions must be chosen carefully, for example in the case of IEX, too high concentrations of salt can promote hydrophobic interactions preventing renaturation, whilst too low salt can promote ionic interactions with the matrix preventing refolding. Folding in this study showed that column refolding is achievable at much higher concentrations than dilution refolding, (compare 5 mg/mL in column refolding compared to 15 μg/mL for dilution refolding). Better success using IEX has been observed with dual gradient systems (Li et al., 2002) where denaturant concentration is decreased whilst salt concentration is increased. Such a system

allows simultaneous elution and refolding as the protein makes its way down the column.

Immobilised metal ion affinity chromatography (IMAC) can be used to bind protein engineered with poly-histidine tails. This specific means of absorption is useful for refolding, and the bonding required is resistant to high concentrations of chaotrope (Li et al., 2004). Refolding is achieved by a gradual decrease in chaotrope concentration, which permits refolding whilst elution is achieved by addition of imidazole (Li et al., 2004). Is it critical in such a system that the poly-His portion does not prevent effective refolding (Li et al., 2004). It is also important that the amounts placed on to the column are not too high, as this may induce aggregation (Li et al., 2004). Metal ions present in these systems can cause oxidation of the refolding protein, which may form incorrectly paired disulfides. Care must also be taken to remove reducing agents, as these may reduce metal ions (Li et al., 2004; Middelberg, 2002). Successful protein refolding by this method has been achieved for a variety of proteins (Glynou et al., 2003; Lemercier et al., 2003; Rehm et al., 2001).

Additionally HIC chromatography can be used. Here its important that the hydrophobic interactions are not be too strong that they prevent refolding (Li et al., 2004). It has successfully been used to refold interferon-α (Guo, 2001), with yields greater than that achieved with dilution or dialysis.

1.4.6.3 Immobilised Catalysts and Molecules Promoting Refolding

An additional form of chromatography is that utilising immobilised folding catalyst such as chaperones (e.g. GroEL), which prevent inappropriate polypeptide interaction by binding to unfolded intermediate forms of proteins. Such immobilised catalysts have been used to refold lysozyme (Dong et al.,

2000). In a separate study, three components of a chaperone system were immobilised on an agarose gel to refold oxidatively scorpion toxin Cn5 (Altamarino et al., 1999). The system comprised: GroEL mini-chaperone (prevents protein aggregation), DsbA (catalyses disulfide-shuffling) and peptidyl-propyl isomerase (catalyse cis-pro isomerisation of proline residues) (Altamarino et al., 1999). In this system it was found that the catalyst could be reused reducing costs considerably. As with any form of affinity chromatography, matrices are expensive to produce and purchase, and more difficult to clean-in-place (CIP) effectively without preventing reuse. Care must be taken when using such systems. Alternatively antibodies can used and may induce their denatured antigen to fold to its native structure (Carlson and Yarmush, 1992).

Catalysts may not necessarily protein based. For example, immobilsed cyclodextrins (Sundari et al., 1999) and immobilised liposomes (Yoshimoto and Kuboi, 1999) could provide effective methods of improving refolding yields.

Immobilsed catalyst methods can provide excellent refolding. Such methods come at considerable cost, where expense is particularly influenced by matrix costs. Such costs may preclude their use at scale (Li et al., 2004).

1.4.7 Novel Methods

A variety of novel methods have been established to improve refolding yields. A particularly interesting development is the use of expanded bed chromatography. Expanded bed chromatography (EBA) combines the binding capability of conventional chromatography, with clarification capacity afforded by the expanded nature of the bed. Such methods can bind target protein from homogenates whilst removing non-target solid phase material such as cell debris.

EBA was used to refold a fusion protein of human growth hormone and glutahione-S-transferase (Cho et al., 2002). IB protein solubilised directly from cell homogenate was bound to the media, cell debris was removed by washing and urea then removed to effect refolding and the folded protein finally eluted (Cho et al., 2002). Use of EBA can reduce the number of renaturation steps required (Cho et al., 2002), improving process economics, but suffers from low capacity.

Another novel means uses synthetic molecular catalysts to refold protein. Machida et al used cycloamylose to successfully refold citrate synthase, carbonic anhydrase B and lysozyme, in solutions containing a variety of ionic and non-ionic detergents (Machida et al., 2000). The inclusion properties of the cycloamylose molecule mean that it can accommodate detergent (which is added to prevent protein aggregation) successfully stripping it from protein structure, promoting correct protein folding (Machida et al., 2000).

An interesting new method is the use of reversed micelles for protein refolding. Reversed micelles are nanostructures, synthesised when surfactant molecules, solvent and small amounts of aqueous solution are combined (Sakono et al., 2004). Sakono et al used reversed micelles to successfully refold ribonuclease A (Sakono et al., 2004). In his method IBs were solubilised and placed into reversed micelles formed by aerosol OT (dioctyl sodium sulfosuccinate). Renaturation was achieved by addition of reversed micelles prepared with redox reagents. Active protein was released from micelles by back extraction into aqueous solvent. This method is effective as it can spatially isolate protein molecules allowing efficient refolding. Addition of the molecular chaperone GroEL improved yields further (Sakono et al., 2004).

1.4.8 An Analysis of Industrial Application of Refolding Techniques

The biotechnology industry requires reliable, consistent, validatable processes that produce product of the highest purity. Every step in a biopharmaceutical process must not only be considered on its ability to give high yields or purity of the target product, but on its throughput and cost. There is no panacea to the refolding of proteins and optimal conditions must often be derived empirically by trial and error. When encountering new proteins considerable effort may be required to refold them to acceptable yields. It is therefore useful to use high-throughput studies to enable optimal refolding. Such studies are most usefully achieved at scale-down level, to save on precious material. Scale-down may well provide clues to how a protein may refold at scale, but it will only effectively do this for very similar processes. Table 1.1 provides a qualitative comparison of the suitability to industrial processes and ability to have perform high-throughput studies. A discussion below details the comparison.

Criterion	Dilution	Dialysis	SEC	Adsorptive Chromatography	Immobilized Catalyst
Ease of use	***	**	**	**	**
Equipment required	***	**	**	**	**
Cost	***	**	**	**	*
Size of equipment	*	**	***	***	***
Concentration refolded at	*	**	*	***	***
Yields	**	**	**	**	***
Use in high throughput studies	***	*	*	*	*
Susceptibility to fouling	***	*	*	*	*
Prevention of aggregation	*	*	***	***	**
Time	*	*	***	***	***

Table 1.1: A qualitative comparison of refolding techniques and their suitability to industrial applications and scale down studies. (*- poor, ** intermediate, *** good,)

1.4.8.1 Ease of Use/ Equipment Required

Dilution requires minimal equipment in the form of a stirred tank and a pump, and is quite easily operated by adding the solution at a predetermined rate by pump, and leaving for a defined period to allow refolding to occur. Automation could probably be achieved relatively simply, but typically such a process would be unlikely to need automation.

Diafiltration will require an ultrafilter and associated ancillary equipment, such as pumps, to circulate flow through the system. Once set up it should be reasonably

easy to operate. Diafiltration can again be achieved relatively easily without automation, but automation may be achieved relatively easily in house.

Chromatography requires a column and ancillary equipment, which may include automated equipment to fractionate and later to visualise chromatograms, and to run pumps. If the process is not automated considerable user interaction will be required to define where stages in the process should begin and end. Many automated chromatography systems are commercially available and this makes the automation of refolding processes relatively easy and preferable especially when using gradients or a variety of buffers

1.4.8.2 Cost

The simplicity of equipment required for a dilution process means that initial costs will be kept at a minimum. It is, however, likely that refolding will take place at low concentration, requiring large tanks, which may come at high cost. Considerable running cost will be incurred in buffer preparation.

An ultrafilter of considerable size will be required for refolding at low concentrations using diafiltration. Running costs will be primarily dictated by the ability for reuse of membranes.

Chromatography costs will initially be taken up by setting up equipment. In a similar fashion to diafiltration, the cost of matrix and its ability for reuse will be critical, especially if fouling by aggregation reduces its reusability.

1.4.8.3 Protein Concentration for Protein Refolding

Protein concentration after refolding is a critical parameter. High concentrations are preferred as this limits concentration required before polishing steps. It has been stated earlier (1.4.4) that for dilution refolding the concentration of protein and refolding yields have an inversely proportional relationship. Greater yields are achieved at lower dilutions, but lower dilutions require tanks of increasing size. Diafiltration methods are affected in a similar way, with greater yields being achieved at low protein concentrations. For dilution and diafiltration applications a compromise must be reached to use concentrations low enough to limit aggregation, but sufficiently high so as to keep tanks/filters to a reasonable size.

Size exclusion chromatography can deal with high concentrations during loading but these undergo considerable dilution upon elution. In some cases this achieves comparable yields and final concentrations of refolded protein to that of dilution refolding.

Adsorptive and immobilised catalyst methods can spatially separate denatured proteins and the dependence on protein concentration is considerably reduced, permitting refolding at higher concentrations.

1.4.8.4 Yields

Good yields can be achieved in refolding by dilution or diafiltration but these are typically achieved at low concentrations. Yields of refolding proteins especially at higher concentrations will be better for chromatographic applications than diafiltration or dilution methods. Immobilised catalyst may achieve especially high yields due to their catalytic activity on the refolding protein.

1.4.8.5 High Throughput Tests

At the simplest level dilution can be achieved at small scale with just a small receptacle such as eppendorf, a pipette to add material and a suitable device to mix to homogeneity. Scale down to microtitre plates is possible for testing a wide a variety of conditions. Using dilution refolding in batch, microtitre plate experiments may give some indication of the performance of chromatographic or diafiltration refolding systems, but they will only reliably predict refolding by dilution at scale.

Chromatography and diafiltration will be difficult to apply at scale down due to the requirement for multiple systems for high throughput tests.

1.4.8.6 Fouling/ Binding

Aggregation may occur in a stirred tank used for refold reactions, and depending on the concentration and degree of mixing a small amount of aggregate, may foul the surface of such vessels, but may be removed with effective CIP. This should not affect the reusability of tanks.

Diafiltration and chromatographic methods will be sensitive to fouling due to their membranes and matrices, and conditions must be more carefully chosen in such methods. Effective CIP may be able to be achieved by caustic and chaotrope washes but this will come at a cost.

1.4.8.7 Prevention of Aggregation

Out of all the methods chromatographic techniques appear to prevent aggregation best. The spatial separation achieved particularly with adsorption methods allows

these methods to prevent aggregation better than dialysis or diafiltration. Such spatial separation permits the use of high protein concentrations, which are not permissible when using dilution and diafiltration-based techniques.

1.4.8.8 Purification Achieved

Dilution and dialysis will only purify the target away from those that don't fold under the conditions used. Dialysis may achieve some removal of contaminants if the molecular weight cut off is appropriately chosen.

Chromatography can achieve the greatest purity after refolding, non-binding species will be eluted, giving considerable purification of samples. In the case of an SEC column the restriction of species of different size may provide for the efficient removal of contaminating species including mis-folded forms.

1.4.8.9 Summary

Clearly chromatographic techniques perform best in terms of throughputs, yields and final concentrations in comparison to the other techniques, especially at scale. The simplicity of dilution refolding, however, makes it highly amenable to high throughput methods. Knowing how various factors, especially those of a chemical nature, affect the refolding process may provide clues to how these factors affect refolding using diafiltration and chromatographic methods. The ease and simplicity of dilution refolding methods influenced the decision to study the effect of factors upon it in this thesis. The following two sections discuss factors, which affect refolding yields, some of which are investigated later in this thesis.

1.4.9 Characteristics of Protein Structure Affecting the Speed of Refolding

The speed of refolding (renaturation) is quantified as the renaturation rate-constant k_r. Gardiner in his PhD study identified the following protein characteristics affecting the renaturation rate (Gardiner, 1988):

- Size of protein
- Number of prolines
- Number of cysteines
- Monomeric or multimeric structure

The larger a protein is, the greater the number of interactions that have to be formed before the protein folds into its native state. Larger proteins are likely to contain greater numbers of domains. The formation of native structure requires domain association, which may become a rate limiting step (Gardiner, 1988).

The half-time of refolding has been shown to be directly proportional to the proline content of the polypeptide (Baldwin and Creighton, 1980). Upon unfolding proline exists in two forms cis (10%-30%) and trans (70-90%) (Brandts et al., 1975). The enthalpy change driving the conversion between the two forms is small. An example of the effect of proline content upon renaturation rates is provided by a study comparing the refolding of two homologous carp parvalbumins, one containing a proline residue the other not (Lin and Brandts, 1978). Measurements of the kinetics of refolding showed that the proline-containing parvalbumin, had an additional slow phase in refolding (Lin and Brandts, 1978). The effect of the number of prolines appears to diminish as

protein size increases, where domain association becomes the rate limiting step (Gardiner, 1988).

As the number of cysteine groups in the protein increases the number of different cysteine pairings that can form increases, this increases the likelihood of incorrect disulfide bonding. The presence of reducing and oxidising agents *in-vitro* prevents this by encouraging disulfide exchange (Gardiner, 1988). Such an exchange may take a significant time before the correct disulfide bonds are formed and hence may limit the rate of renaturation.

After refolding of individual subunits, an additional step of subunit association is required for oligiomeric proteins (Gardiner, 1988). The rate of renaturation is not limited by this association, but is believed to be limited by conformational changes. Refolding of multimeric proteins has been shown to have approximately second order kinetics (Rudolph et al., 1976; Rudolph et al., 1977a). Refolding to a quaternary structure is not absolutely necessary for protein activity, but is essential for full functionality (Gardiner, 1988). Final protein refolding concentration is important in such systems. The concentration of subunits where aggregation is minimised is generally lower than those used for monomeric proteins. It is believed this is the case as the subunits possess a greater amount of surface available to interact with other subunits (Gardiner, 1988). A study of the refolding of pig heart lactate dehydrogenase (Rudolph et al., 1977b) showed that yield was reduced at both high and very low protein concentrations. Yields at high concentrations were reduced because of incorrect molecular interactions, whilst yields at lower concentrations were thought to occur because of reduced subunit interactions.

1.4.10 Process Factors Affecting Refolding

Apart from the protein itself, conditions used for both solubilization and within the refolding reactor can determine how quickly denatured proteins can be renatured.

The low refolding yields at high protein concentration has been well documented (De-Bernardez Clark et al., 1998; Goldberg et al., 1991; Yasuda et al., 1998). Refolding is a monomolecular reaction, involving intramolecular reactions and can be described by first-order kinetics, whilst aggregation is multi-molecular and is described by higher order kinetics (De-Bernardez Clark et al., 1998). Given the competition and order of the reactions the formation of native protein is favoured at low protein concentrations (Buswell et al., 2002).

According to Gardiner one the most important factors affecting the renaturation rate is the conformation of the protein after IB solubilization (Gardiner, 1988). If full denaturation is achieved the protein will be left as a randomly coiled polypeptide whose residues are exposed to the solvent environment and whose disulfides are all reduced. In practice, it may be preferable to subject the protein to partial denaturation as this may permit faster refolding times. Conservation of protein structure during solubilization is particularly advantageous for large complex proteins, which have slow renaturation rates. An example of partial denaturation increasing renaturation rates, when compared to full denaturation, was demonstrated in the refolding of prochymosin (Gardiner, 1988).

It is important to keep temperature at ambient or near ambient to prevent structural destabilisation (Gardiner, 1988). Raising temperature will increase the

energy of the system, which may lead to an increase in collision frequency and aggregation. High temperatures can also induce protein denaturation and are best avoided. The effect of temperature upon protein refolding is aptly demonstrated by the folding bovine muscle pyruvate kinase. The ability of bovine muscle pyruvate kinase to renature decreases above 32°C, but remains constant between 0°C and 25°C (Porter and Cardenas, 1980). The decrease in yield was associated with aggregation and incorrectly formed monomers.

Temperature also affects molecular interactions in the protein itself (Cleland and Wang, 1993). Hydrophobic interactions, a major driving force for folding, are endothermic (Dill, 1990). This means that an increase in temperature will increase the rate at which these bonds are formed (Cleland and Wang, 1993). This may improve the rate of refolding of the protein, but hydrophobic interactions also have a major influence on aggregation, which could be detrimental to yields.

Denaturants, such as urea and GdHCl have been shown to bind directly to proteins (Cleland and Wang, 1993). As the bulk concentration of the denaturant decreases so does that associated with the protein allowing the protein to refold (Cleland and Wang, 1993). During the reduction in denaturant concentration the protein will form stable intermediates. The rate of formation of these and the fully folded native structure is dependent on rate of removal of denaturant (Cleland and Wang, 1993). If removal is achieved too quickly incorrect folding and aggregation could result. Removal too slowly however, could lead to intermolecular association resulting in aggregation (Gardiner, 1988). Such effects are incredibly protein-specific.

The final denaturant concentration can have significant effects on refolding yields. Refolding of tryptophanase from 8M urea to an intermediate denaturant concentration caused aggregation (London et al., 1974). An expanded conformation is thought to have allowed interactions to occur resulting in aggregation. Despite this, non-denaturing concentrations of chaotrope (as high as 1.3 M GdHCL) have been shown to have positive effects on the refolding of a variety of proteins including hen-egg white lysozyme (Hevehan and De-Bernardez Clark, 1997), carbonic anhydrase II (Wetlaufer and Xie, 1995), *P. fluorescens* lipase (Ahn et al., 1997) and interferon β polypeptides (Dorin et al., 1996). Denaturant concentrations that are too low, can also cause aggregation. This effect was apparent for the refolding of chymotrypsinogen where over-dilution resulted in aggregation (Orsini and Goldberg, 1978).

The presence of stabilising salts in the refolding buffer will increase the rate of renaturation. Such salts form unfavourable interactions with the hydrophobic moieties of the polypeptide, forcing these moieties to associate with each other to limit the area exposed (Gardiner, 1988). This effect increases with salt concentration. Addition of stabilising salt can induce the formation of native structure but high concentrations may also cause aggregation. High ionic strengths have been shown to result in aggregation of γ-interferon when present in the refolding buffer (Hsu and Arakawa, 1985).

pH is particularly important to cysteine side-chains which have a pKa of approximately 8.5 (Stryer, 1995). Epstein and Goldberger revealed in studies on the renaturation of lysozyme that the rate of formation of correct disulfides

(evident by yield of active protein) was accelerated at pH close to the pKa of cysteine (Epstein and Goldberger, 1963). At more acidic pHs (e.g. pH 7) the thiol of cysteine is more stable and less likely to form disulfide bonds (Gardiner, 1988).

High alkalinity has been used with success to refold prochymosin (Kawaguchi et al., 1984; Marston et al., 1984). Refolding at pH 10-10.7 improved refolding yield. It is likely that this effect is not only due to the reactivity of cysteine residues at this pH, but also because high pH maintains a loose protein structure and decreases intermolecular interactions by charge repulsion (Gardiner, 1988).

When choosing pH, isoelectric conditions should be avoided so as to minimise the likelihood of aggregation. The pH should be at the upper limit of the pH for stability for the native structure of the protein and high enough to accelerate disulfide exchange (Rudolph et al., 1997).

The redox environment for protein refolding is critical for enabling efficient disulfide formation and shuffling, and hence formation of the native structure.

Redox ratio is a critical factor in oxido-shuffling systems utilising both reducing and oxidising species. Optimal ratios vary and must typically be found empirically. Literature suggest that generally a 10:1 ratio of reducing to oxidising species is sufficient (Ahmed et al., 1975). The reducing environment is required to reduce the stability of mixed disulfides, encouraging native protein-protein disulfide bonds. Previous studies suggest that for lysozyme a redox of ratio from 3:1 (reduced: oxidised) to 1:1, and total thiol concentrations of 5-15 mM produce the greatest yields (Hevehan and De-Bernardez Clark, 1997).

As discussed earlier, air oxidation provides a means of disulfide shuffling. Disulfide shuffling systems utilising air are limited by the slow mass transfer of

rate oxygen into solution (De-Bernardez Clark, 2001). Increased rates of agitation will improve mass transfer, but may lead to aggregation due to increased shear and interfacial stress (De-Bernardez Clark, 2001). A comparison of air oxidation and oxido shuffling (glutathione mixture) systems was conducted with ribonuclease A (Ahmed et al., 1975). It was found that protein concentrations of 0.025 mg/mL were inhibiting for renaturation using an air oxidation system, whilst there was no concentration dependence for refolding using the oxido-shuffling systems, below 0.125 mg/mL (Gardiner, 1988). It therefore appears that oxido-shuffling system although more expensive (due to cost of redox reagents), should be chosen to get the best refolding yields.

Mixing, especially that at the macroscale will lead to concentration gradients across the system. Of particular importance to refolding reactions are the concentrations of denaturant and protein. The concentration of denaturant has a direct impact upon the flexibility and solubility of the protein. The rate of decrease in denaturant concentration influences the ability of the protein to refold correctly. The greater the concentration of protein, the greater the likelihood that collisions of partially folded protein will occur, leading to an increase in the degree of aggregation. Mixing must therefore provide efficient dispersion to prevent aggregation.

The rate of mixing can have a considerable impact on the rate of chemical reactions. The speed of refolding is a strong function of the protein studied. In some cases, e.g. the Arc repressor, refolding may take microseconds (Renzeperis et al., 1999), whilst some proteins take many years to refold, e.g. α- lytic protease (in the absence of its amino-terminal pro region) (Sohl et al., 1998). Given the requirement for process efficiency in industry, targets for refolding processes

have an upper limit of the order of hours. In the majority of situations it seems unlikely that mixing will affect the rate of the refolding reaction *per se*. However it may affect the rate of aggregation and hence alter the selectivity between native soluble protein formation and aggregation.

The effect of salt and denaturant on refolding reactions has been discussed above but other additives can have considerable effect upon protein refolding yields. Typically additives will operate by stabilising the native state, by specifically destabilising non-native folding forms, or by increasing the solubility of refolding intermediates (De-Bernardez Clark, 1998). It is believed additives achieve binding non-specifically or by preferential hydration of the protein resulting in an increase in the number of water molecules at the protein surface (Cleland and Wang, 1993).

A variety of reagents have been shown to improve refolding yields. Sugars such glycerol and glucose have been shown to form more compact hydrated protein structures (Arakawa and Timasheff, 1982). Studies have shown that refolding of beta-lactamase was improved by addition of glucose (Valax and Georgiou, 1991). Addition of sugars does not necessarily produce positive effects; ribonuclease refolding was slowed down in the presence of glycerol (Tsong, 1982). Other examples of sugars improving refolding yields include: Hen Egg White lysozyme HEWL (Glucose) (Maeda et al., 1996), and *P.fluorescens* lipase (Glycerol) (Ahn et al., 1997).

Some refolding additives operate by binding non-specifically to the protein, altering the reaction pathway (Cleland and Wang, 1993). An example of this are detergents, which will improve refolding by preventing aggregation (De-

Bernardez Clark, 1998). Effects of detergents are indicated in reviews (Cleland and Wang, 1993; De-Bernardez Clark, 1998). Positive effects have been observed for several proteins: rhodonase (lauryl maltoside) (Tandon and Horowitz, 1987), human growth hormone (Tween) (Bam et al., 1996), carbonic anhydrase (Triton-X-100) (Wetlaufer and Xie, 1995), and interferon β-polypeptides (SDS) (Dorin et al., 1996). A problem with detergents are they are difficult to remove because of their ability to form micelles and to bind to protein (De-Bernardez Clark, 1998). To combat this an interesting method was developed (Rozema and Gellman, 1996), in which protein is first exposed to detergent (to prevent aggregation). The detergent is then stripped from the protein using cyclodextrin. This system has been termed "artificial chaperone-assisted refolding" and has been used to refold carbonic anhydrase (Rozema and Gellman, 1996).

Polyethylene glycols (PEG) are useful refolding additives, and are less hydrophobic than detergents (Cleland and Wang, 1993). It was shown that PEG weakly binds to the first molten globule intermediate in the carbonic anhydrase B pathway (Cleland and Randolph, 1992). PEG did not however increase the rate of refolding, but aggregation was considerably inhibited (Cleland et al., 1992b; Cleland and Wang, 1990a). It has been postulated that a PEG–protein intermediate complex forms that is incapable of association, this then folds to form native protein (Cleland et al., 1992b). PEG has also been shown to improve yields of human deoxyribonuclease and human plasminogen activator (Cleland et al., 1992a).

The amino acid L-arginine has been shown to be an excellent refolding additive improving refolding of HEWL (Hevehan and De-Bernardez Clark, 1997), Fab-fragment (Buchner and Rudolph, 1991) and alpha-glucosidase (Stempfer et al., 1996). A study showed that acetone, acetoamide or a urea derivative improved the refolding of lysozyme (Yasuda et al., 1998) showing the importance of chemical structure on determining the effect on refolding.

A recent study has revealed the effect of a variety of refolding additives upon refolding kinetics (Dong et al., 2004). Refolding additives comprise two groups, the first of which includes acetamide, acetone, thiourea and L-arginine, which stabilise unfolded protein and intermediates. Kinetically they reduce the refolding and aggregation rate with increases in concentration (Dong et al., 2004). There is an optimal concentration for high refolding yields. The second group of folding additives, which includes glycerol stabilise protein structure. With these additives both refolding rate and yield are increased (Dong et al., 2004).

Inefficiencies in IB processing steps (1.2.4) mean that IB are typically contaminated with cell debris when solubilised and hence these contaminants will be present in the refolding reaction. Cell debris also consists of the insoluble matter from within the cells i.e. cell walls, membranes etc. IB preparations are typically contaminated with peptidoglycans, membrane proteins, lipid and nucleic acids (Thatcher, 1990). Contaminating cell debris has been shown to increase the degree of proteolysis of IB (presumably through the presence of proteolytic enzymes) (Wong et al., 1996), and affect the yield from refolding

(Georgiou and Valax, 1999; Maachupalli-Reddy et al., 1997). Therefore efficiency of purification steps may be critical to refolding yield.

It has been shown previously that refolding yields can be improved by decreasing the rate of injection of the denatured protein (Katoh et al., 1999; Katoh and Katoh, 2000). In the case of lysozyme, the formation of native protein is favoured at low protein concentrations (Buswell et al., 2002) where aggregation of folding intermediates is slower than refolding to the native protein. Similarly, the gradual addition of denatured protein to refolding buffer results in improved yields due to the presence of a lower concentration of partially refolded protein over the duration of the addition (Katoh et al., 1999).

The time required for the refold is a critical factor especially when considering application in industry. Evidently the longer the renaturation time the greater the yields. However renaturation times should also consider protein stability, and therefore limits may have to placed on renaturation times.

1.5 Lysozyme as a Model System

The first two chapters of this thesis use lysozyme to study the effect of process factors upon refolding yield. Several papers have investigated the mechanism of lysozyme refolding (Buswell and Middelberg, 2002; Buswell and Middelberg, 2003; Kiefhaber, 1995; Radford et al., 1992; Roux et al., 1997; Wildegger and Kiefhaber, 1997) using different methods to characterise the steps involved. An appreciation of the mechanisms of lysozyme refolding may permit the

determination of how factors can affect refolding yields and makes this a good model protein.

1.5.1 Lysozyme Structure and Function

Lysozyme is a glycosidase, it hydrolyses the [[beta]]-1,4, glycosidic linkages between N-acetylmuramic acid and N-acetyl glucosamine, which are found in the cell walls of gram-positive bacteria, promoting the lysis of such organisms (Stryer, 1995). Lysozyme is found in both animal and plant kingdoms, as well as some viruses. It is found in mammals in bodily secretions, where it augments the response of secreted immunoglobulin A acting as a humeral antibiotic. In addition it has been found in several plant species including papaya and ficus.

One of the most commonly used proteins in refolding studies is hen egg-white lysozyme (HEWL) (De-Bernardez Clark et al., 1998; Katoh et al., 1999; Lee et al., 2002; Maachupalli-Reddy et al., 1997; Yasuda et al., 1998). Lysozyme from egg white has a molecular weight of 14.4 kDa (Sophianopoulos et al., 1962). It contains 4 disulfide bonds. The tertiary structure is a two lobe globular structure separated by a cleft across one side of the molecule, which contains an active site that can accommodate six units of the substrate 2-acetamido-2-deoxy-D-glucopyranose (Branden and Tooze, 1999). The first lobe contains five α-helices, the second is dominated by a three-stranded anti-parallel beta pleated sheet (Branden and Tooze, 1999). A structural model of lysozyme is detailed in the Figure 1.13.

Figure 1.13: Structure of Hen Egg White Lysozyme. 6lyz.pdb (Diamond, 1974)

1.5.2 Lysozyme Mechanism

The mechanism of action of lysozyme has been described (Kirby, 2001) (Figure 1.14). The reaction occurs with the retention of configuration; the water group, which replaces the N-acetylglucosyl group, attaches to the same face of the sugar. The reaction occurs by a double displacement. The carboxyl group of Asp 52 acts as a nucleophile to form the glycosyl enzyme intermediate in a SN2 type reaction. The enzyme carboxylate is then displaced from the glycosyl-enzyme intermediate by water, completing the reaction and placing the water in the same conformation as the N-acetyl-glucosamyl group it replaced.

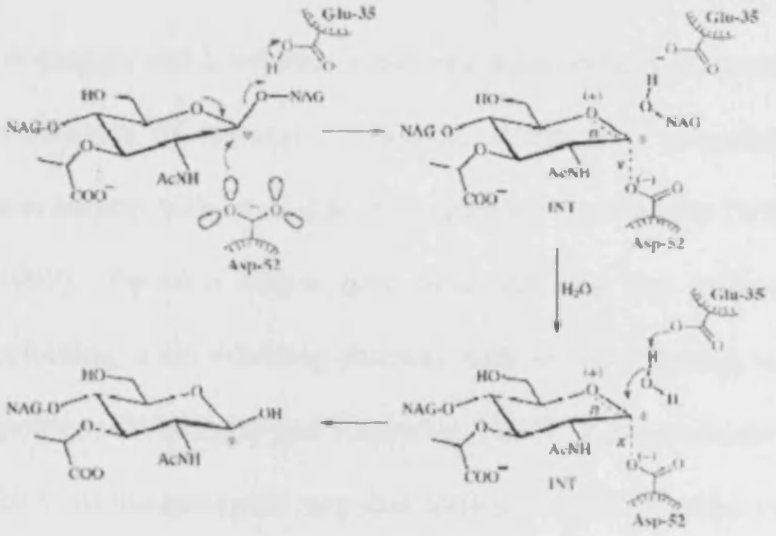


Figure 1.14: Mechanism of action of lysozyme (Kirby, 2001)

1.5.3 Proposed Mechanism(s) of Lysozyme Refolding

Lysozyme refolding is a complex kinetic process that has been investigated by various groups. It is thought to comprises three steps:

1. Hydrophobic collapse occurring early in the refolding process (Wildegger and Kiefhaber, 1997).
2. A well-defined intermediate state, which is formed with a relaxation time (1/k) of 35 ms (Wildegger and Kiefhaber, 1997). Studies of the structure of this intermediate were achieved through a combination of 2D NMR and pulsed hydrogen exchange (Radford et al., 1992), and mass spectrometry (Miranker et al., 1993). These studies revealed that the α-domain (consisting entirely of α-helical structure) is already folded in the intermediate but the β-domain consisting of primarily β-strands is still amorphous (Wildegger and Kiefhaber, 1997).
3. Conversion of the intermediate to native protein (1/k=365 ms) (Wildegger and Kiefhaber, 1997).

Wildegger and Kiefhaber conducted a particularly interesting investigation of the mechanics of lysozyme refolding, where they proposed a triangular folding mechanism with an energetically trapped intermediate (Wildegger and Kiefhaber, 1997). Previous studies have concluded that two pathways exist for lysozyme refolding, a slow-folding pathway with an intermediate, and an alternative direct pathway (Wildegger and Kiefhaber, 1997). It has been shown that approximately 20 % of the lysozyme may fold through this direct pathway (Kiefhaber, 1995). It is perplexing why the slow pathway exists. A description of the refolding pathways proposed are described in Figure 1.15A and B.

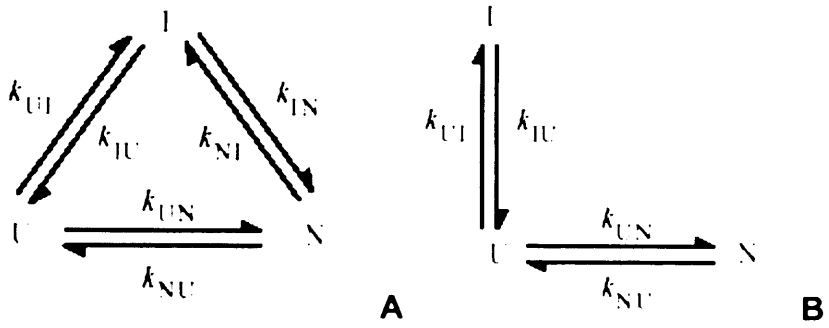


Figure 1.15: A schematic representation of the mechanisms proposed for lysozyme refolding. A: triangular mechanism, B: Dead end mechanism. kXY: rate constant for conversion between forms. U:unfolded form, I = intermediate. N: Native form. (Taken from (Wildegger and Kiefhaber, 1997)

The first mechanism above is the triangular mechanism. The second is the dead-end mechanism, and assumes that the intermediate is on a slow unproductive pathway. Renaturation by this route would require unfolding first to the unfolded state and then conversion to the native state (Wildegger and Kiefhaber, 1997). To distinguish between these models, the dependence of the refolding reaction of guanadinium chloride (GdmCl) was assessed. Experiments of guanidine dependencies in the presence of sodium sulphate (which stabilises the intermediate structure) revealed that the dependencies of the two observable rate constants agreed with the triangular mechanism and contradicted the result expected if the dead end mechanism was true (Wildegger and Kiefhaber, 1997).

A slightly different mechanism has been proposed by examining the kinetics of recovery of native structure after refolding by both activity and circular

dichroism. The following models were proposed (Figure 1.16) (Roux et al., 1997):

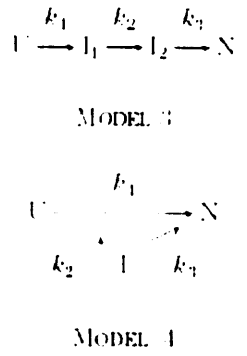


Figure 1.16: Models proposed by Roux et al for lysozyme refolding (Roux et al., 1997) U: unfolded form, I_1: Intermediate 1, I_2: Intermediate 2, N: Native structure, k_n: rate constant for reaction

Model 4 was discounted as the value obtained for k_3 was far too close to 0 (i.e. the reaction would be far too slow) (Wildegger and Kiefhaber, 1997). Results fit with model 3 where I_1 is a species that has recovered 20% of its native structure, whereas I_2 is a partially active form of the enzyme, which has recovered the ellipicity of the native state at the wavelength used in circular dichroism. The conversion of I_2 to native structure requires minor changes over a considerable time ($k=0.0002\ min^{-1}$), not detectable by circular dichroism. This model confirms observations from other experiments (Sawano et al., 1992) which showed that three out of four possible three-disulfide derivatives of HEWL had enzymatic activities 70-80% of the native form, suggesting that I_2 may just be an incompletely oxidised form of HEWL with only three out of the four disulfides formed (Roux et al., 1997).

1.5.4 Refolding Versus Aggregation in the Lysozyme System

The kinetic properties of the productive (refolding) and abortive (aggregation) pathway for lysozyme has been assayed (Goldberg et al., 1991). It was proposed that there was a stage in the refolding where a protein becomes irreversibly committed to refolding and a stage at which it becomes irreversibly committed to aggregation. The majority of aggregation occurred in the first five seconds, whereas the commitment to renaturation was much slower occurring at 4.5 minutes (Goldberg et al., 1991).

Further studies have revealed that the kinetics of lysozyme aggregation are almost independent of starting concentration, with aggregation levelling off at the same point despite a 2-fold difference in concentrations at the most extreme cases (Goldberg et al., 1991). These results suggested that commitment to renaturation happens before 1 minute hence preventing any further aggregation, and contradicting the results obtained earlier where commitment to renaturation was proposed to occur in 4.5 minutes. To explain these contradicting phenomena the following model was proposed (Figure 1.17). It is important to note at this point that the alternative fast pathway proposed by Kiefhaber where the renaturation does occur via a transient intermediate had not been discovered (Kiefhaber, 1995). This faster alternative pathway might describe why approximately 20% of protein refolds, even at high concentrations of denatured protein. As such, this pathway may even out-compete the aggregation pathway. The model proposed is detailed in Figure 1.17:

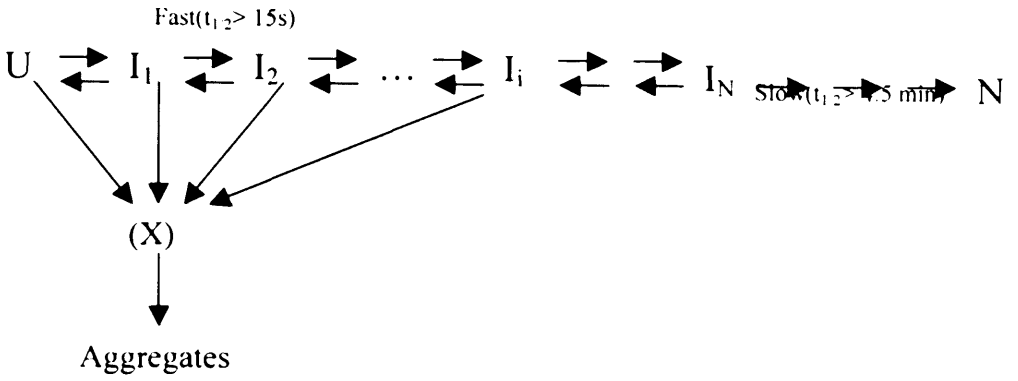


Figure 1.17: Proposed model for lysozyme refolding inclusive of aggregation. U: unfolded form, I_n: Intermediate, N: native, X: intermediate multimer (Goldberg et al., 1991).

This model proposes that the reaction goes through a series of intermediates before the native state is formed. All intermediates up to I_i can form aggregates (possibly via intermediate x), whereas none of the steps after I_i can lead to aggregation. Steps between I_1 and I_N occur rapidly and are reversible, whereas those beyond I_N are slow and irreversible. If the transition to I_N is rapid, i.e. takes approximately 1 minute, then after a minute the levels of aggregatable intermediate may be too low for aggregation to occur, therefore, allowing refolding. It is tempting to suggest that the intermediates in this pathway may correspond to various semi-folded states of the protein (Goldberg et al., 1991).

Studies of the kinetics of egg-white lysozyme refolding have been reported in the literature which reconsider aggregation (Buswell and Middelberg, 2002). The original model of refolding assumes that it is a monomeric first order reaction that

competes with the higher order aggregation reaction. These studies have revealed that this model is an oversimplified view of refolding kinetics. Addition of labelled native lysozyme during refolding experiments showed that the label was clearly incorporated into aggregates. This suggests that a competing reaction is occurring whereby native protein can be incorporated into aggregates. This is achieved either via an intermediate or by direct surface incorporation. Current models do not account for such reactions and will, therefore, over-estimate yields of refolded proteins (Buswell and Middelberg, 2002).

Although the models presented in the literature review are somewhat contradicting, it is evident that aggregation is a multimeric reaction and will proceed faster than the slower monomeric refolding reaction. One object of this thesis is to understand how process factors such as mixing can affect the competition between refolding and aggregation, and hence final refolding yields.

1.6 Project Aims

The literature review has provided a grounding in refolding, as well as introducing IBs and the steps required prior to refolding from them.

It is the primary aim of this thesis to study the effect of process factors upon protein refolding yields. The introduction has given details of factors which affect protein refolding. Considerable work has been undertaken to establish the effect of these factors, but areas still remain to be explored and it is the aim of this thesis to understand some of these.

Numerous studies have sought to optimise the chemical environment within the refolding reactor (De-Bernardez Clark et al., 1998; Hevehan and De-Bernardez Clark, 1997; Yasuda et al., 1998). Very few however have studied the effect of

physical processes (Buswell et al., 2002; Lee et al., 2002) on tryspinogen and lysozyme respectively. Perhaps the most important of these is mixing. Papers have established that there is an effect of mixing (Buswell et al., 2002; Lee et al., 2002) but have failed to analyse critically why. The first part of this thesis studies why mixing affects protein-refolding yields.

The second part of this thesis will examine the interactions between factors affecting protein-refolding yields. To date such analyses have looked at factors in isolation, without considering factors in combination. A factorial study will be undertaken to understand interactions between chemical and physical process parameters. Such an understanding can be critical to determining how refolding methods can be optimised.

The first two sections of this study use pure protein systems to understand the effect of process factors upon refolding yield. Since IBs are typically contaminated after isolation, the final parts of this thesis examines the efficacy of inclusion body purification steps and seeks to understand how contaminant levels affect and influence eventual refolding yield.

2 A Critical Assessment of the Impact of Mixing on Dilution Refolding.

2.1 Introduction

2.1.1 Refolding as a Process Bottleneck

The refolding of proteins from inclusion bodies often presents a bottleneck in the generation of recombinant proteins expressed in *E.coli* (Buswell and Middelberg, 2003; De-Bernardez Clark, 1998). Yields at industrially relevant concentrations are restricted by aggregation of protein upon dilution of the denatured form (De-Bernardez Clark et al., 1998; Goldberg et al., 1991; Lee et al., 2002). Numerous studies have sought to optimise the chemical environment within the refolding reactor (De-Bernardez Clark et al., 1998; Hevehan and De-Bernardez Clark, 1997; Yasuda et al., 1998) very few however have studied the effect of physical processes (Buswell et al., 2002; Lee et al., 2002). Perhaps the most important of these is mixing.

The following section reviews the definition of mixing and the processes that contribute to the degree of mixing in a system. This is followed by an analysis of the likely impact of mixing events on protein refolding.

2.1.2 An Introduction to the Fundamentals of Mixing

Mixing creates a homogenous fluid environment where components of the fluid mixture are dispersed. Mixing is never instantaneous; the time required for "complete" mixing of a reactor is its mixing time. This may be defined as the minimum time to achieve a predetermined degree of homogeneity from a completely segregated state (Doran, 1995). Mixing time is frequently utilised to

characterise mixing within stirred tanks, despite its inherent inability to characterise mixing locally in favour of characterising the whole reactor (Guillard and Trägårdh, 2003). In well-designed reactors, with a turbulent flow regime, the time required for mixing is of the order of 0.1 to 10 seconds (Bourne, 1985). Complex system dynamics can be quantitatively or even qualitatively affected by mixing, when the mixing time is comparable to the characteristic reaction time (Fang and Lee, 2001). Such dynamics may affect yields from refold steps.

The following processes may qualitatively describe the mixing in a reactor (Oldshue, 1983):

Distributive mixing: Relatively large eddies exchange positions, and convect material so that uniformity at a macroscopic level size is established (Bourne, 1985). This typically occurs in the order of seconds.

Dispersive mixing: The large eddies described above are broken up by turbulent shear and a mixture of finer grade is created (Bourne, 1985). The mixture however remains segregated at the molecular level (Bourne, 1985).

Diffusive mixing: In the finely dispersed structure, diffusion occurs over small distances to randomise the mixture at the molecular scale, producing a homogenous fluid (Bourne, 1985). This typically occurs in the order of milliseconds.

Although these mixing processes have been described in a consecutive manner it is more likely that they occur simultaneously (Buswell et al., 2002). Mixing phenomena throughout this chapter shall be described using these qualitative descriptions.

In the case of stirred tank reactor (STR), distributive mixing is characterised as the fluid motion created by the pumping action of an impeller circulating liquid within the reactor. Such flow creates a concentration distribution within the reactor, which influences both the reaction rate and selectivity in situations where multiple reactions occur (Oldshue, 1983).

The limit of homogeneity achieved by dispersive mixing is defined by the smallest eddies that can be formed in a particular fluid under defined conditions (Doran, 1995). This may be given by the Kolmogorov scale of turbulence λ. Its size is directly proportional to the kinematic viscosity of the fluid being pumped, and is indirectly proportional to the energy dissipation rate, which at steady state is equal to the power input of the impeller. At the level of these eddies there is little mixing, and attainment of homogeneity depends upon diffusion.

Diffusional mixing, is mixing on a microscale where segregation and molecular diffusion become important (Oldshue, 1983). Diffusion is the only mechanism that permits contact between molecules and is therefore a necessary precursor to chemical reactions (Bourne, 1985). Diffusional mixing time is highly dependent on the size of the eddies. The larger they are the longer it will take material to diffuse across them. This mixing time also depends on the molecular diffusivity, which is a function of the materials that are being mixed and the temperature. The greater the diffusivity the shorter time it will take for a mixture of molecules to become homogenous at a molecular scale. Mixing does not affect the inherent rate of diffusion, but since energy dissipation affects the size of eddies it will affect the contribution made by diffusional mixing.

2.1.3 Effect of Mixing on Chemical Reactions

The rate of mixing can have a considerable impact on the rate of chemical reactions. If the rate of reaction significantly exceeds the rate of mixing, a region close to the point where reagents contact each other will exist where concentrations deviate from those corresponding to complete homogeneity, and a situation of segregation is said to exist (Oldshue, 1983). Under these conditions rates of reaction are determined by the diffusion of the reactants to the reaction site. Such a reaction is said to be instantaneous, and will therefore be affected by the contribution of diffusive mixing, which in turn is dependent on the distributive and dispersive mixing. If a chemical reaction occurs at a rate significantly slower than the mixing time, mixing will be over before the reaction has had a chance to proceed to any significant extent (Bourne, 1985). In this case segregation is absent and the inherent rate of reaction is determined by chemical kinetics (Oldshue, 1983). Between these states, i.e. where mixing and reaction rate are more comparable, a situation exists where the rate of reaction is controlled by both chemical kinetics and the level of mixing.

2.1.4 Effect of Mixing on Refolding

Mixing, especially that at the macroscale will lead to concentration gradients across the system. Of particular importance to refolding reactions are the concentrations of denaturant and protein. The concentration of denaturant has a direct impact upon the flexibility and solubility of the protein. The rate of decrease in concentration influences the ability of the protein to refold correctly. The greater the concentration of protein, the greater the likelihood that collisions of partially folded protein will occur, leading to an increase in the degree of aggregation. The speed of refolding is a strong function of the protein studied. In

some cases, e.g. the Arc repressor, refolding may take microseconds (Renzeperis et al., 1999), whilst some proteins take many years to refold, e.g. α- lytic protease (in the absence of its amino-terminal pro region) (Sohl et al., 1998). Given the requirement for process efficiency in industry, targets for refolding processes have an upper limit of the order of hours. In the majority of situations it seems unlikely that mixing will affect the rate of the refolding reaction *per se*. However, it may affect the rate of aggregation and hence alter the selectivity between native soluble protein formation and aggregation.

2.1.5 Lysozyme as a Model Protein

This Chapter looks at the effect of mixing upon the refolding of lysozyme. Lysozyme refolding may be described by a simplified pathway, where native lysozyme forms from the unfolded state via an intermediate (Figure 2.1). Lysozyme was chosen as a suitable test protein, as numerous studies have been undertaken on its refolding (Buswell and Middelberg, 2002; Buswell and Middelberg, 2003; Kiefhaber, 1995; Radford et al., 1992; Roux et al., 1997; Wildegger and Kiefhaber, 1997). The refolding and aggregation reactions compete for a folding intermediate. Refolding is a monomolecular reaction, involving intramolecular reactions and can be described by first-order kinetics, whilst aggregation is multi-molecular and is described by higher order kinetics (De-Bernardez Clark et al., 1998). Given the competition and order of the reactions the formation of native protein is favoured at low protein concentrations (Buswell et al., 2002). The refolding of lysozyme typically takes between minutes and hours, depending on conditions, whilst formation of intermediates occurs between 10 and 200 milliseconds (De-Bernardez Clark et al., 1998).

Given the higher order of aggregation reactions it is expected that under high final protein concentrations (>0.1 mg/mL), suitable for industrial application, aggregation will be faster than refolding. This assumption has been confirmed in studies (De-Bernardez Clark et al., 1998).

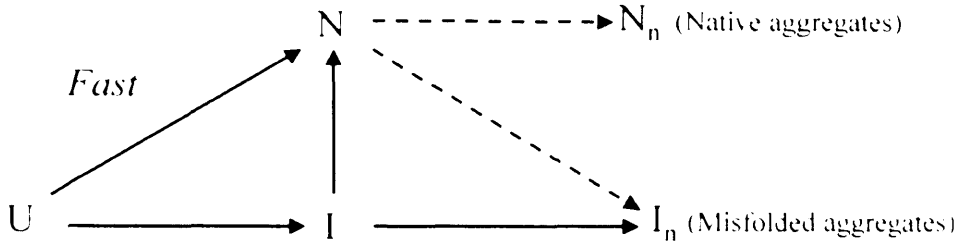


Figure 2.1: A simplified schematic of the refolding pathway for lysozyme. Lysozyme can refold by two pathways primarily via an intermediate, but evidence suggests that lysozyme can also refold via an alternative fast pathway which does not go via the intermediate (Kiefhaber, 1995). Intermediate can polymerise to form irreversible aggregates that cannot be converted to native form. Evidence from recent studies suggests that native protein may bind to aggregates (Buswell and Middelberg, 2002). Native aggregates only occur at very high protein concentrations or under conditions utilised in crystallisation procedures. Key N: native protein, U: Unfolded protein, I: Intermediate, I_n: Aggregate, N_n: Native aggregate.

Work has shown that the redox ratio is critical to refolding yields (De-Bernardez Clark et al., 1998; Hevehan and De-Bernardez Clark, 1997). Refolding in this study utilises a refolding buffer that contains only the oxidised form of the redox reagent. Excess reducing agent in the denatured protein solution added to the refolding solution provides reducing power to form the redox pair (cystamine: cysteamine) over time.

Evidence from various groups suggests that protein concentration is inversely proportional to the yields obtained from refolding reactions (De-Bernardez Clark et al., 1998; Goldberg et al., 1991; Yasuda et al., 1998). Various other groups have found that slower addition of the denatured protein to refolding buffer gives improved yields (Katoh et al., 1999; Katoh and Katoh, 2000). Under such slow injection rates, it is likely that the influence of macroscopic concentration gradients will be minimised even under poor mixing conditions. As such any effect of mixing may be described by efficiencies in dispersive mixing, which in turn directly influences the efficiency of diffusive mixing. A previous study (Lee et al., 2002) showed that mixing intensity in such a system improved refolding yields, but did not attempt to link this with the conditions of mixing.

2.1.6 Chapter Aims

The objective of this chapter is to investigate the linkage between energy dissipation levels and protein refolding yields. In a stirred tank reactor (STR) the maximum energy dissipation achievable is defined by the maximum impeller speed available. This is not only limited by the capacity of the motor but also by the need to prevent the formation of air-liquid interfaces, which have been shown to cause precipitation of native protein (Maa and Hsu, 1997) and may have deleterious effects upon the selectivity of refolding over its competing aggregation reaction. Given these constraints a two-impeller system was designed to provide increased local energy dissipation rates close to the point of reagent addition, and avoid the high shear effects in the proximity of large bulk impellers operating at high velocity so as to achieve effective bulk mixing. The design includes a small paddle impeller (mini-impeller) close to the injection

point. It was envisaged that the additional impeller would increase the energy dissipation rate experienced by the injected material upon entry into the reactor in an efficient manner with the consequence that the overall levels of energy used to achieve good contacting within the vessel would be reduced. Knowledge of the effect of energy dissipation upon protein refolding yields may influence reactor design so as to maximise protein-refolding yields.

2.2 Materials and Methods

2.2.1 Materials

Chicken egg white lysozyme, lyophilised powder, ~50,000 units/mg protein was purchased from Sigma-Aldrich Company Ltd (Poole, Dorset, England, UK). All additional chemicals except those listed below were sourced from Sigma-Aldrich Company Ltd (Poole, Dorset, England, UK) and were of at least reagent grade quality.

Sulfuric acid, potassium phosphate dibasic (K_2HPO_4) and potassium dihydrogen phosphate tri-hydrate ($KH_2PO_4{\cdot}3H_2O$) were purchased from BDH Limited, VMR International Ltd (Poole Dorset, England, UK). Phenolphthalein in methylated spirits, was purchased from Fisher Scientific UK (Loughborough, Leicestershire UK).

2.2.2 Equipment Utilised

The bulk impeller was a Heidolph R2R-2000, (Heidolph Instruments, Schwabach, Germany). A Eurostar digital, IKA® Labortechnick stirrer (Staufen, Germany)

drove the mini-impeller. Spectrophotometric measurements were performed on a Genesys6, from Thermo-Spectronic (Rochester, NY, USA). Temperature was measured using a 2000 series temperature probe, from Jencons (Scientific) (Leighton Buzzard, Bedfordshire, UK). A water bath used to keep the reactor at defined temperature (25° for all experiments) (SU6, Grant Instruments Ltd, Cambridge UK).

2.2.3 Stirred Reactor

The experiments used a twin-impeller reactor (figure 2.2) of 0.25 L capacity with heating cooling jacket: D_T= 63 mm, H_T= 80 mm. The tank was baffled so as to increase the efficiency of mixing, 4 baffles were arranged as shown in Figure 2.2 (w= 7 mm, D_D= 19 mm, W_B= 6mm, H_B = 4.5 mm. The flat-blade Rushton impeller was used to allow comparison with previous studies on the effect of mixing on refolding (Buswell et al., 2002; Lee et al., 2002). The injection point (d= 0.75mm) was placed at angle of 70° and at a distance of 2 mm from the mini-impeller blade. The mini-impeller was a 2 blade paddle impeller: D_I = 10mm W_B= 6mm, H_B =2 mm. This "mini-impeller" was placed at approximately 85° to the horizontal, 16 mm above the bulk impeller blade, 5 mm off centre from the bulk impeller, and 2 mm below the injection point.

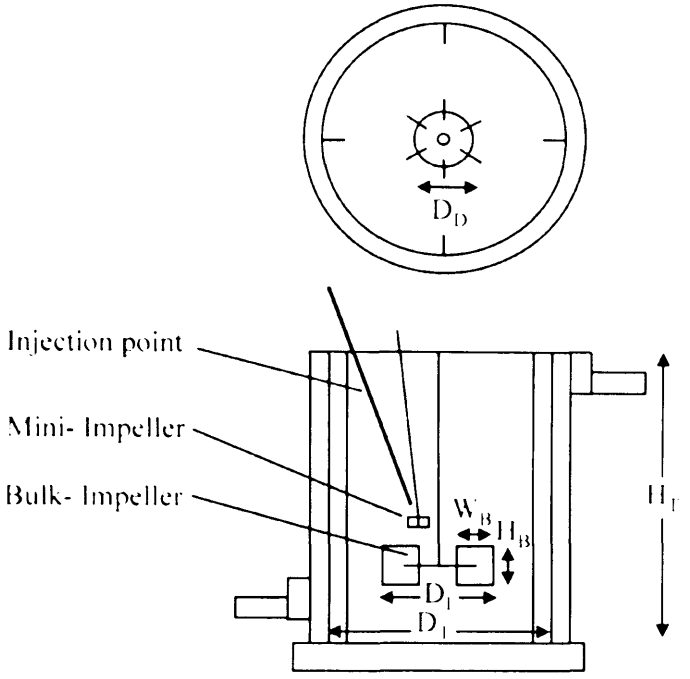


Figure 2.2: Scale Down Twin-Impeller reactor. The reactor has two impellers the first the bulk impeller which controls bulk flow within the reactor. The second a 2 blade paddle positioned just below the injection point to disperse rapidly materials upon injection into the reactor. Reactor configuration. Tank: D_T= 63mm, H_T= 80 mm Bulk Impeller: Six Blade Rushton impeller D_I= 24 mm, D_D= 19 mm, W_B= 6mm, H_B =4.5 mm. Height in tank = 21 mm. Mini-impeller: 2 Blade paddle D_I = 10mm W_B= 6mm, H_B = 2 mm. at approximately 85° to the horizontal, 16 mm above bulk impeller blade, 5mm off centre from bulk impeller, 2 mm below the injection point. Injection point approximately at 70° to vertical, 30 mm from bulk impeller

2.2.4 Reynolds Number as an Assessment of Mixing Intensity

Mixing intensities were quantified as Reynolds numbers for ease of comparison with other work and were calculated according to the following formula:

$$Re_i = \frac{N_i D_i^2 \rho}{\mu} \qquad (2.1)$$

Where Re_i= Reynolds number for an impeller, N_i = Impeller speed (s^{-1}), ρ = density (kgm^{-3}) , D_i= impeller diameter (m), μ= viscosity (Pa.s). The viscosity and density of the refolding solution will vary over the time course and between different addition rates, it is however likely that these will not vary by more than 20% and in the majority of instances by far less. The Reynolds numbers calculated therefore present an approximation of the average

2.2.5 Assessment of Critical Injection Time and Effect of Impeller Intensity Upon Micromixidness Ratio.

The parallel competing reaction system described by Fournier and Falk in 1996 (Fournier et al., 1996) was used as a means to assess the efficiency of dispersive mixing and therefore the contribution of diffusive mixing in the reactor.

The chemical basis of the reaction system is the coupling of borate neutralisation (2.2) and the Dushman reaction (2.3) (Lin and Lee, 1997):

$$H_2BO_3^- + H^+ \leftrightarrows H_3BO_3 \text{ (Quasi-Instantaneous)} \qquad (2.2)$$

$$5I^- + IO_3^- + 6H^+ \leftrightarrows 3I_2 + 3H_2O \text{ (Fast)} \qquad (2.3)$$

The reactions are conducted with acid as the limiting reagent; the two reactions therefore compete with each other for protons. If the slower Dushman reaction is allowed to occur, the iodine formed may further react with iodide ions to yield tri-iodide ions:

$$I^- + I_2 \leftrightarrows I_3^- \qquad (2.4)$$

If diffusive mixing is efficient reaction (2.3) will not have time to occur and the majority, if not all, of the protons available will be taken by the borate reaction (Lin and Lee, 1997). In the case of inefficient diffusive mixing a proportion of protons will go through the Dushman reaction producing iodine and subsequently tri-iodide ions. The concentration of these may be measured spectrophotometrically. Samples are measured at 353 nm using an extinction coefficient of 2395.9 m^2/mol (Guichardon and Falk, 2000). Diffusive mixing efficiency may be characterised by the segregation index. The segregation index , x_s, is a number from 0 to 1, which gives the proportion of protons going to make iodine as a proportion of the total available (Guichardon and Falk, 2000). In the case of total segregation all the protons go through the Dushman reaction, and the value of x_s is 1. When diffusive mixing is instantaneous no segregation is present and the value of x_s is 0. The value of x_s was calculated using equation (2.5) (Guichardon and Falk, 2000).

$$x_s = \frac{Y}{Y_{SI}} \qquad (2.5)$$

Where Y is the ratio of the number of moles of acid that go through reaction (2.3) to the total number of moles of acid injected and is calculated from formula (2.6) (Guichardon and Falk, 2000):

$$Y = \frac{2(n_{I_2} + n_{I_3^-})}{nH^+} \qquad (2.6)$$

The mole number n of tri-iodide is derived from the measured tri-iodide levels. Iodine mole number is calculated from a mass balance of the iodine atoms, and the mole number of acid is calculated from the amount of acid injected. Y_{SI} is the value of Y when total segregation occurs, that is when all the acid injected goes through reaction (2.3) (Guichardon and Falk, 2000).

The segregation index may also be used to generate the micromixidness ratio, α, which represents the volume fraction of reactor contents that are perfectly mixed at the microscale versus the volume fraction that stays completely segregated (Assirelli et al., 2002).

$$\alpha = \frac{1 - X_s}{X_s} \qquad (2.7)$$

The segregation index is influenced by injection rate. For a fixed volume the rate of injection is inversely proportional to the injection time. At shorter injection times the reagent plume injected will not be well-dispersed, and higher segregation indices will result through the influence of distributive mixing. As the injection time increases the value of the segregation index becomes constant. The point at which injection time no longer influences the segregation index is defined as the critical injection time. Beyond the critical injection time segregation index is only influenced by the efficiency of diffusive mixing, which is a function of the energy dissipation rate and dispersive efficiency (Guichardon and Falk, 2000).

The effect of injection time upon the segregation index was measured at the lowest bulk impeller intensity used (Re_b = 576). H_2SO_4 (0.2 M, 2 mL) was injected into the reaction solution (200 mL of a solution containing 0.1818 M H_3BO_3, 0.0909M NaOH, 2.3 mM KIO_3, 0.011167 M KI, mixed in that order) at flow rates corresponding to injection times of 24 to 375 seconds. A P500 syringe pump (GE Healthcare, Uppsala, Sweden) was used to achieve a smooth addition rate. On completion of addition the reaction mixture was mixed for two minutes and then samples were taken. The segregation index was calculated according to equation (2.5) as previously described (Guichardon and Falk, 2000) and was used

to infer the critical injection time. Each injection was repeated in triplicate for accuracy. The average error by standard deviation of the measured segregation index was ± 3.6 %.

2.2.6 Effect of Impeller Speed Upon Segregation Index

The effect of bulk-impeller intensity upon segregation index was determined at a flow rate, which delivers the acid in the critical injection time. H_2SO_4 (0.2 M, 2 mL) was injected into the reaction solution (described above) at 1.3 mL/min and at bulk impeller intensities corresponding to Re_b 959 - 3836 and mixed for 2 min on completion of acid addition. Segregation index and micromixidness ratio were calculated as described before. Error of α was an average of ± 4.6% (1 standard deviation).

2.2.7 Effect of Bulk and Mini-impeller Mixing Intensity Upon Segregation Index

The effect of bulk and mini-impeller mixing intensity upon segregation index was determined at various combinations of bulk and mini-impeller intensities H_2SO_4 (0.2 M, 2 mL) was added at 1.3 mL/min to a 200 mL aliquot of the reaction solution. The solution was mixed for a further 2 minutes at the selected combination of mixing intensities to ensure the reaction went to completion. Segregation index and micromixidness ratio were calculated as described before. Error by standard deviation was an average of +/- 3% (1 standard deviation).

2.2.8 Assessment of Mixing Time in the Twin Impeller System

Mixing time was measured by the use of a neutralization reaction tracked by the indicator phenolphthalein. Water (200 mL), NaOH (1 M, 2 mL), phenolphthalein (1% in methylated spirits (0.25 mL) were mixed to form a deep pink solution. HCl (1M, 2 mL) was added to the top of the alkali solution at a variety of bulk and mini-impeller intensities quantified as Reynolds numbers Re_b and Re_m respectively. The time from the end of addition to the solution going colourless was taken as the mixing time for the reactor at the combinations of Re_b and Re_m used. Each combination of Re_b and Re_m was repeated in triplicate, with an average standard error of +/- 7% (1 standard deviation) in terms of the observed mixing time.

2.2.9 Denaturation-reduction of Lysozyme

Lysozyme was denatured and reduced in denaturing buffer (4.5 M GdHCl, 32 mM DTT, 1 mM EDTA, 50 mM Tris @ pH 8) for 2h at a concentration of between 9.28-9.67 mg/mL. The concentrations of denatured protein were assessed after 1:40 dilution in acetic acid (0.1 M) using an extinction coefficient of 2.37 ml cm^{-1} mg $protein^{-1}$ (De-Bernardez Clark et al., 1998; Lee et al., 2002). The denatured state was confirmed by *Micrococcus lysodeikticus* lysozyme assay and RP-HPLC as described in the subsequent two sections.

2.2.10 Micrococcus Lysodeikticus Lysozyme Assay

The denatured state was confirmed using the method of Lee et. al. (Lee et al., 2002). A 100 µL aliquot of the 1:40 diluted denatured protein solution in acetic acid was diluted in 2.9 mL of *Micrococcus lysodeikticus* cell suspension (0.3

mg mL in 0.1 M potassium phosphate buffer at pH 7). The change in absorbance at 450 nm was tracked over 1.5 minutes, after a delay of 30s to allow equilibration. Absence of enzymic activity was taken as a confirmation of the denatured state.

2.2.11 RP-HPLC Assay for the Assessment of Native Lysozyme Concentration

Assessment of native lysozyme concentration was achieved by an HPLC method (Lee et al., 2002). A C_5 reversed phase column (5 μm, 300 Å, 150 mm x 4.6 mm, Jupiter column (Phenomenex, Macclesfield UK)) was used on a Beckman Gold High Pressure Liquid Chromatography system (High Wycombe, UK) comprising a 126 Pump Unit, 166 Detector unit, 507e Autosampler unit, using System GOLD software. A linear acetonitrile water gradient 0.1 TFA (v/v) (30-46 % over 12 minutes) at a flow rate of 1 mL min was used. 200 μL samples were applied to the column. A standard curve was created using native lysozyme and used to calculate the yield of refolded lysozyme in samples. Native lysozyme concentration was assessed at 280nm using an extinction coefficient of 2.63 ml cm^{-1} mg^{-1} (De-Bernardez Clark et al., 1998). Native lysozyme had a retention time of 8.64 ± 0.08 min, whilst denatured protein had a retention time of 11.24 min ± 0.01 min.

2.2.12 Assessment of the Effect of Mixing Duration Upon Lysozyme Refolding Yield

Denatured lysozyme at 9.32 mg mL was renatured by 20-fold dilution of denatured lysozyme into refold solution (5 mM Cystamine, 1 mM EDTA, 100 mM Tris pH 8). Addition was achieved by peristaltic pump (Pharmacia P1, GE

Healthcare, Uppsala, Sweden) at a rate of 1.19 mL/min (504 s) at various bulk and mini-impeller intensities. Mixing for 2h was then applied at the range of mixing intensities chosen. 2 mL samples were removed at 5, 15, 30, 60 and 120 minutes and left to stand for 24h at ambient temperature. Yields of renatured lysozyme at 24 hours were then determined by RP-HPLC for the various mixing durations and intensities.

2.2.13 Assessment of the Effect of Impeller Intensity Upon Refolding Yield

Denatured lysozyme at 9.28 mg/mL was renatured by 20-fold dilution into the refold solution. Addition was achieved as described in the previous section. The refold mixture was then mixed for 5 minutes at the bulk and mini-impeller intensities chosen. Samples were taken and left for 24h at ambient temperature. Yields of renatured lysozyme after 24 hours were measured by RP-HPLC. Samples were performed in duplicate with a maximum error of +/- 5% (standard deviation).

2.2.14 Assessment of the Effect of Mixing and Injection Time on Lysozyme Refolding Yield

Denatured lysozyme at 9.67 mg/mL was renatured by 20-fold dilution into refold solution at addition rates of 1.19 mL/min, 0.115 mL/min or rapidly by pipette (> 30 ml/min) at the same feed position. For each experiment mixing was conducted for a total of 2h (inclusive of injection time) at the bulk and mini-impeller intensities chosen for the experiment. Samples were then taken and left for 24h at ambient temperature. Yields of renatured lysozyme after 24 hours were measured by RP-HPLC. Runs for an addition rate of 1.19 mL/min were performed in

duplicate with an average error of +/- 3.6% (standard deviation). These were used as an approximation of error for the other injection speeds.

2.3 Results and Discussion

2.3.1 Assessment of Critical Injection Time

It can be seen in Figure 2.3 that as expected the segregation index for the iodate-iodide reaction decreased with injection time. The critical injection time under the worst mixing conditions used (Re_b= 576) appears to be approximately 90s. Beyond this point, it is expected that macroscopic concentration gradients are minimised, and injection time has no further effect on segregation index. The effect of impeller speed upon diffusive mixing efficiency was evaluated under these injection conditions.

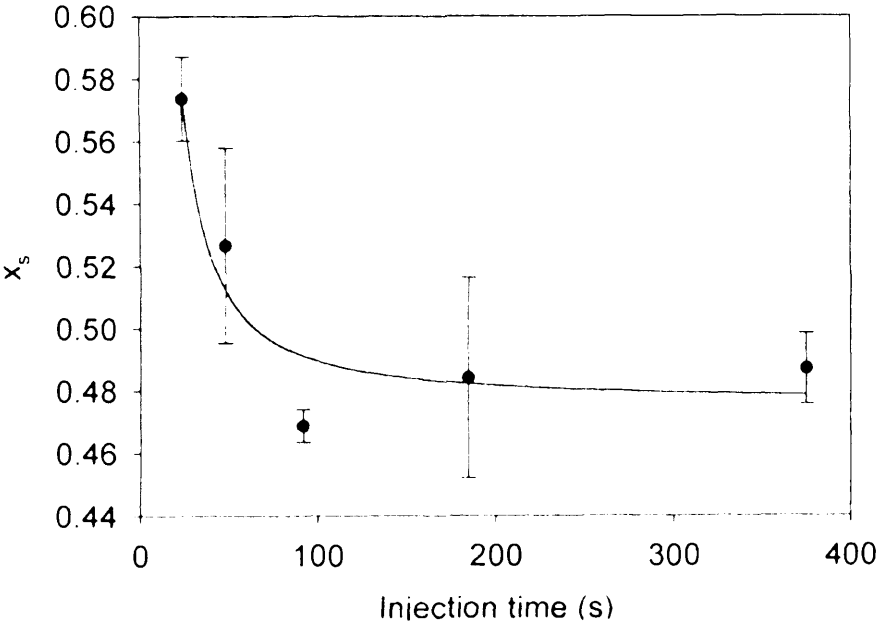


Figure 2.3: The effect of injection time upon segregation index (x_s) for the iodide/iodate reaction.

2.3.2 Characterisation of Dispersive Mixing in the Stirred Tank Reactor

The extent of dispersive mixing was assessed under a range of bulk impeller speeds. Results are reported in terms of the micromixidness ratio (defined by equation (2.7)). It can be seen from Figure 2.4 that an increase in bulk impeller speed increases the micromixidness ratio. This may be explained as follows. The lifetime of any fluid droplet and its size is determined by the energy input, which is a direct function of the impeller speed (N). The greater the value of N, the more efficient the dispersive mixing and the smaller the droplet. Given that the rate of borate neutralisation is faster than the Dushman reaction, hydrogen ions (H^+) are preferentially used in the former. The preference for the borate neutralisation reaction means that a set amount of H^+ will go through this route. Smaller droplet sizes have shorter lifetimes, and hence a reduced amount of dispersed acid will be available to go through the Dushman reaction. This then results in a smaller value of segregation index, which is related to α by equation (2.7) leading to the trend in Figure 2.4. The linearity in the response suggests that at the chosen injection rate the assay is only measuring the contribution of diffusive mixing. If this were not the case the relationship, especially at low impeller intensities and segregation index, would be non-linear. Any non-linearity would result from significantly decreased micromixidness ratios at impeller speeds where the rate of injection is faster than that defined by the critical injection time. However, it is likely that as Re_b increases further, a limit will be reached where the linearity is lost. Typical refold reactors tend to operate at low Re_b, so this condition may not be reached.

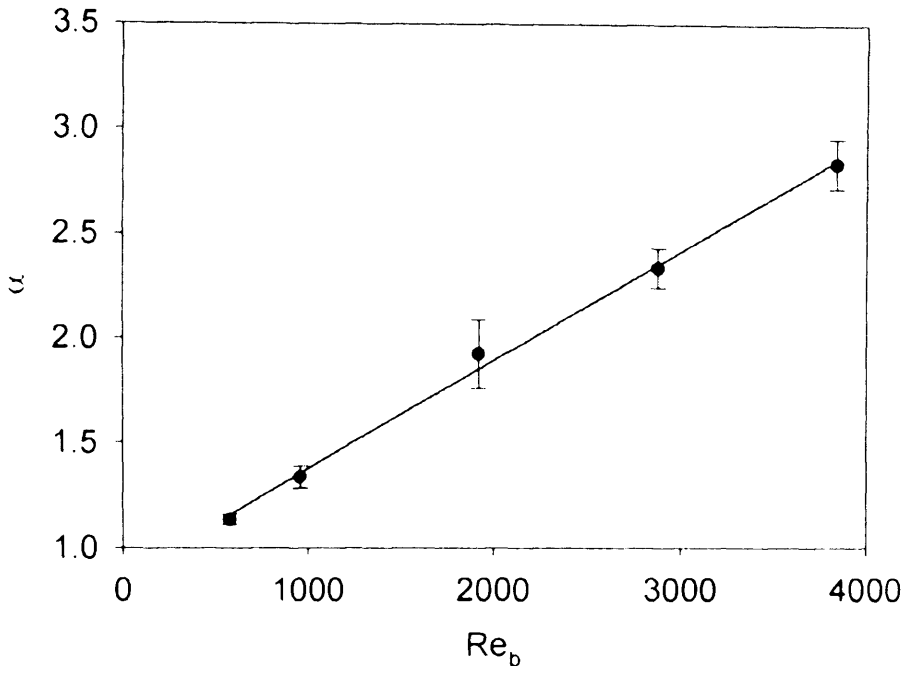


Figure 2.4: The effect of bulk impeller intensity (Re_b) upon micromixidness ratio (α) for the iodate-iodide reaction at t_{inj}= 92s.

According to turbulence theory, greater energy dissipation levels lead to a more significant contribution of diffusive mixing. This prediction has been experimentally verified (Fournier et al., 1996). It has also been subsequently confirmed by showing a directly proportional relationship between micromixidness ratio α and local energy dissipation rate (Assirelli et al., 2002). Assessment of the individual effects of the mini-impeller on the dispersive mixing cannot be accurately quantified given the complex mixing present in the twin impeller system. Data using the parallel competing reaction system (Figure 2.5) however, clearly indicate that the mini-impeller has a considerable effect upon the segregation index. It appears that the effect of the bulk impeller is reduced in the presence of the mini-impeller. The mini-impeller has a greater influence on micromixing on the whole due to its greater proximity to the injection point.

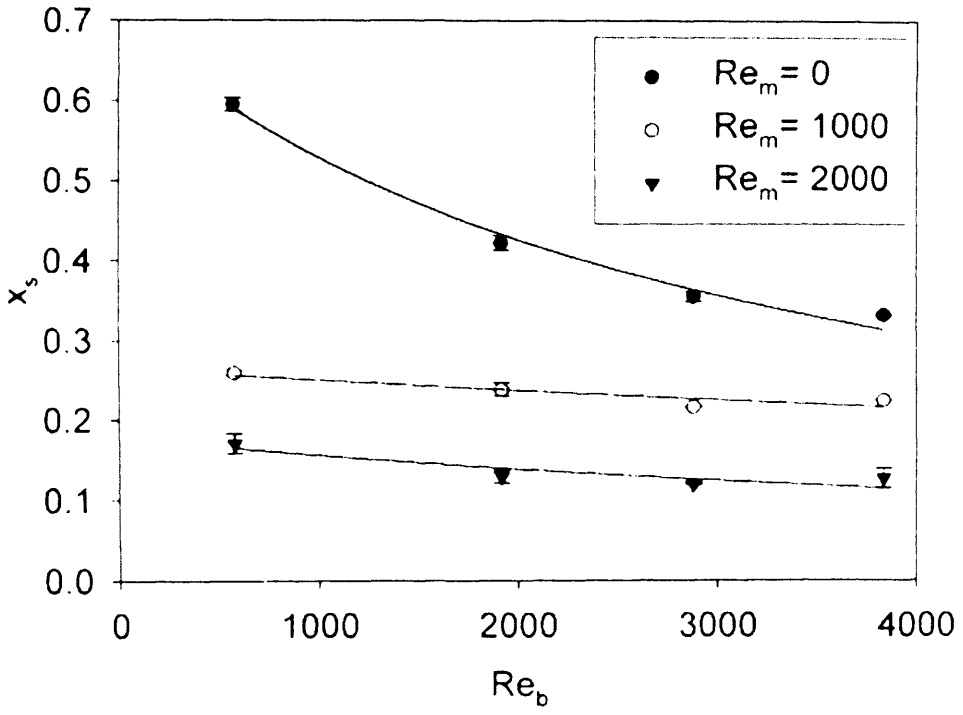


Figure 2.5: Effect of bulk impeller intensity (Re_b) and mini-impeller intensity (Re_m), on segregation index.

2.3.3 Effect of Mini-impeller Upon Mixing Time

To ascertain whether the effect of the mini-impeller was limited to just that of increasing the rate of dispersive mixing, distributive mixing was characterised under a variety of bulk and mini-impeller intensities. It is clear from the data in Figure 2.6 that the mini-impeller has only a limited effect upon mixing time. The effect of the mini-impeller appears to be most apparent where bulk mixing is minimal. A comparison of the relative effects of the two impellers reveals a far greater effect of the bulk impeller upon distributive mixing. This relationship is expected given that the bulk impeller is larger and is located in a position that has been proven to increase mixing efficiencies (Shiue and Wong, 1984). Mixing

times appear to correlate well with data from a previous studies (Hoogendoorn and den Hartog, 1967; Shiue and Wong, 1984).

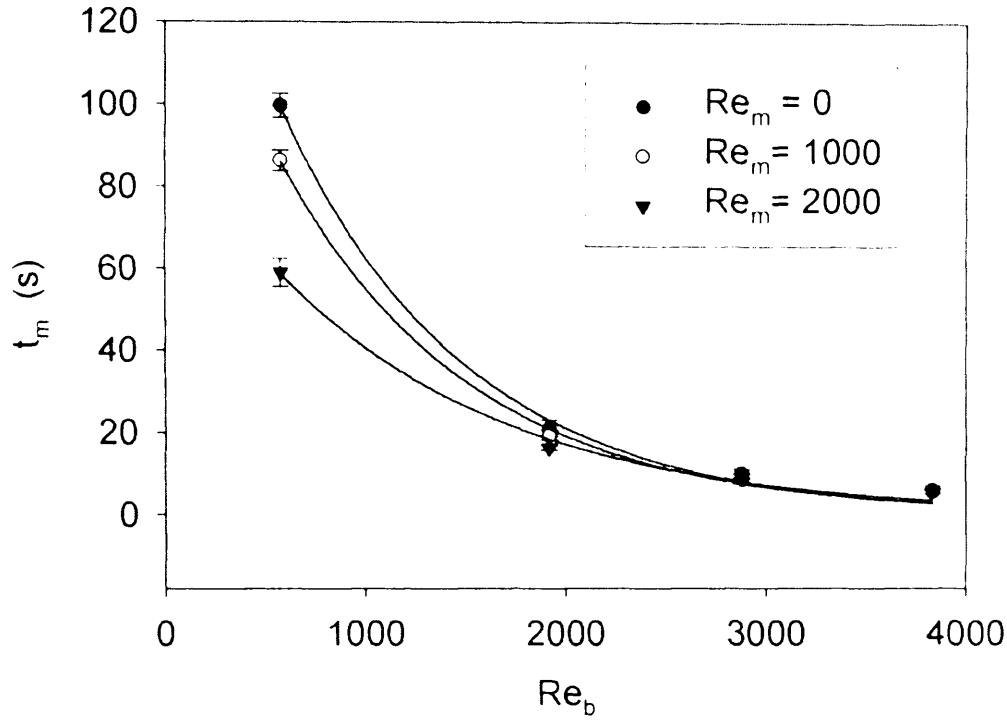


Figure 2.6: Effect of bulk (Re_b) and mini-impeller (Re_m) intensities upon mixing time (t_m) for the phenolphthalein-tracked neutralisation reaction.

2.3.4 Lysozyme Refolding in the Twin Impeller System

The aim of this chapter was to examine the effect of mixing upon final refolding yield. For experimental efficiency it was important to establish the earliest point in time at which mixing as complete. It can be seen from Figure 2.7 that mixing was essentially complete under all conditions in less than 2 minutes. Therefore to ensure homogeneity a conservative practical minimum of 5 minutes was set. It can be seen in Figure 2.7, that extending the mixing beyond 5 minutes had no significant effect upon refolding yield. This suggests that the initial dispersion of the protein is key to refolding yields, and that mixing beyond the attainment of homogeneity has only a limited effect upon refolding. Evidently mixing beyond this point confers little selectivity over the competing aggregation and refolding

reactions. Further to this it appears that beyond 5 minutes the protein molecules have already segregated according to whether they will aggregate or refold. This observation agrees with earlier work (Goldberg et al., 1991) which, for lysozyme, suggested that commitment to aggregation occurs over a half-time ($t^{1}/_{2}$) of less than 30 seconds (0.185- 1.48 mg/mL), whilst commitment to refolding had a half time ($t^{1}/_{2}$) of 4.5 minutes (0.1 mg/mL). The rate of reaction for refolding is therefore slower than the mixing time and mixing will not affect the refolding rate itself, as mixing will be complete before the refolding reaction has had a chance to proceed to any significant extent (Bourne, 1985). Mixing may, however, affect the rate of aggregation, as this proceeds at a rate closer to that of the mixing time.

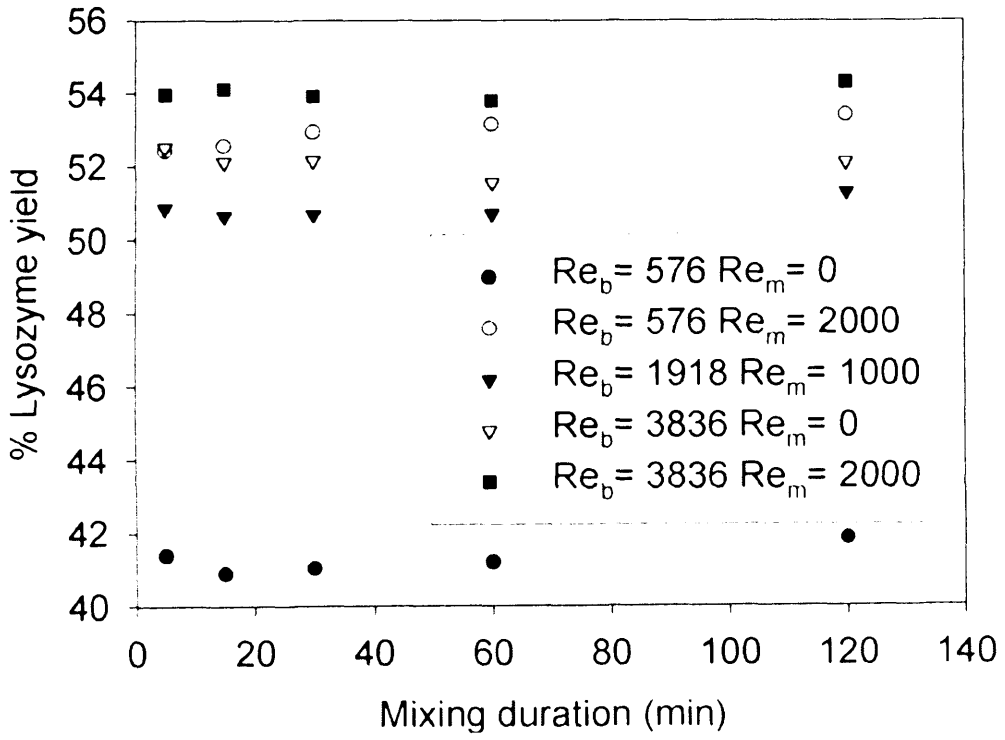


Figure 2.7: Effect of mixing duration and intensity upon refolding yields of lysozyme at Q= 1.19mL/min.

2.3.5 Effect of Mixing Conditions Upon Lysozyme Refolding Yield

After establishing the effect of mixing duration upon refolding yields, the effect of mixing intensity upon refolding was studied. It can be seen in Figure 2.8, that mixing intensity increased the refolding yields, but only to a limited extent. The effect of bulk-impeller intensity appears to be the most significant at the intensities used, with the degree of refolding showing a similar pattern to the micromixidness ratio. It is clear from the data that the main effect of the mini-impeller is only really apparent when the bulk impeller speeds are low. The mini-impeller located close to the injection point does have an affect on protein refolding yield, and this appears to be a function of the local energy dissipation rates created. The results suggest that the local energy dissipation rate may affect refolding yields. The effects of this additional energy dissipation are minimal at high bulk impeller speeds and it is likely that a critical energy dissipation rate exists, beyond which further increases have minimal effect. It may therefore be inferred that the efficiency of dispersive mixing (which in turn affects the efficiency of the diffusive mixing process) is critical to lysozyme refolding yields. For mixing to influence the yield of a reaction, the reaction must proceed at rate similar to the mixing time. In this system the refolding reaction is too slow to be affected by mixing and it is more likely that it is the aggregation reaction or commitment to the aggregation reaction (early steps that form intermediates as in Figure 2.1) that are affected by the level of mixing.

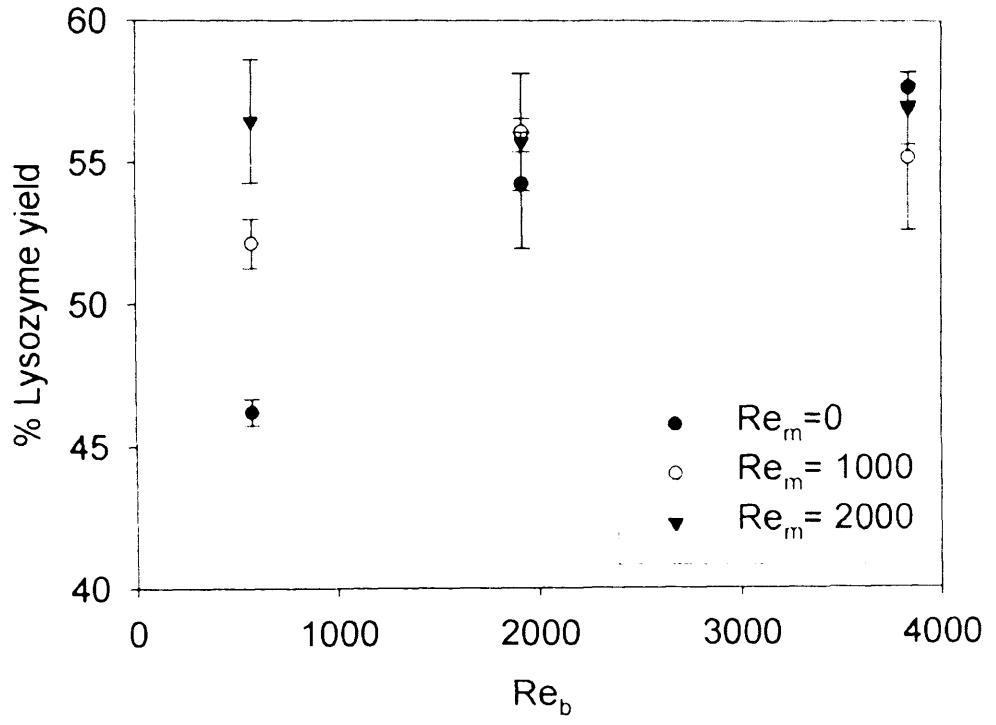


Figure 2.8: Effect of bulk (Re_b) and mini-impellers (Re_m) upon refolding yield. Q= 1.19 mL/min.

2.3.6 Effect of Injection Rate and Mixing Upon Lysozyme Refolding Yields

A further facet of protein refolding is how quickly the protein is added to the refolding buffer. The results from experiments to study the importance of injection rate are detailed in Figure 2.9. The slower the rate of protein addition, the greater the yield of refolded protein. At slower injection rates refolding will occur at a lower concentration at any particular time. Lower concentrations serve to reduce aggregation and increase yields. Inefficiencies at all scales of mixing and the rapid establishment of the final protein concentration, results in lower yields for refolding under batch conditions. Under the slowest injection rate used (a flow rate corresponding to an injection time that is approximately ten times the critical injection time), an effect of mixing was still observed. The observation of an effect of mixing extends considerably beyond the critical injection time, and

confirms the effect of dispersive mixing on refolding yields. The effect of the mini-impeller appears to increase with injection time. This is probably an effect of the increased exposure of the injected material to the local energy dissipation of the mini-impeller. At high flow rates the mini-impeller may be flooded with injected material and will not be able to convey its full effect. The slower the flow-rate the more likely it is for the denatured protein to be affected by the energy dissipation rates imparted by the mini-impeller where mixing efficiency is improved. This further strengthens its effect upon refolding. The dispersive mixing achieved by the mini-impeller may provide sufficient dispersion of protein molecules to prevent aggregation, increasing the selectivity of the refolding reaction over that of aggregation.

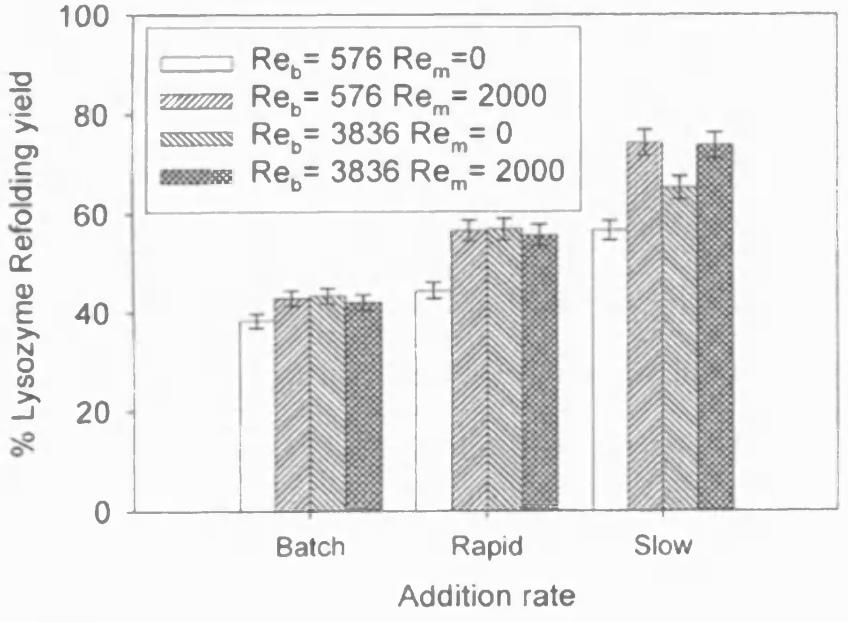


Figure 2.9: Effect of bulk impeller mixing intensity (Re_b), mini-impeller intensity (Re_b), and addition rate (batch >30 mL/min, rapid 1.19 ml/min (injection time just over critical injection time), slow 0.115 ml/min (injection time approximately 10x longer than critical injection time) upon refolding yields in the twin-impeller reactor.

2.4 Conclusions

This chapter characterises dispersive mixing in a scale-down twin impeller reactor, and seeks to relate mixing efficiency with refolding yields. The effect of distributive mixing upon the segregation index appears to be minimal beyond 90s reagent injection time. Increases in impeller speed decreased the segregation index, by virtue of the greater levels of energy dissipation rates afforded at the higher impeller speeds. Assessment of the effect of the mini-impeller, revealed that it had only a minimal effect upon distributive mixing.

Analysis of the effect of various mixing conditions within the reactor revealed that at injection rates below the critical injection rate (rate which delivers the defined volume in the critical injection time), refolding was affected by impeller speed, confirming that dispersive mixing, and therefore energy dissipation affects refolding yields. The mini-impeller appeared to affect refolding yield at this injection rate, but its effect was limited to low bulk impeller speeds. It is proposed that the rate of aggregation is affected by inefficiencies in dispersive mixing. Such inefficiencies increase the likelihood of protein molecule collisions and increase aggregation to the detriment of refolding yields. This chapter considers only a single set of impeller geometries. It would however be of interest to examine the effect of other impeller geometries and positions upon refolding yield, but this was beyond the scope of the current work.

As expected protein refolding yield increases with injection time. Under batch addition conditions all levels of mixing are sub-optimal, high protein concentrations prevail and yields are poor. Yields increase with a decrease in injection rate, in part due to the increases in distributive mixing efficiency. At the

slowest injection rate studied (approximately 10 times slower than the critical injection rate), mixing still had an effect upon refolding yields. This confirms that dispersive mixing affects refolding yields. The effect of the mini-impeller increased as the rate of injection fell, presumably since the injected material has greater levels of exposure to the local energy dissipation created by the mini-impeller.

It is clear from the experiments that levels of the energy dissipation experienced by the injected material is critical to refolding yields, and suggests a possible role of dispersive mixing upon refolding yield.

The second chapter has provided an explanation for the effect of mixing upon refolding. Mixing, however is just one of the factors that can affect protein refolding. To gain a greater appreciation of the effect of process factors it is important to understand how a variety of process factors interact to affect refolding yields. To achieve this Chapter 3 details a factorial experiment to understand interaction between process factors and their effect upon protein refolding yields.

3 Factors Affecting Protein-refolding Yields in a Fed-batch and Batch-refolding System.

3.1 Introduction

Expression of protein as inclusion bodies offers several processing advantages over expression as soluble protein. Inclusion bodies are resistant to proteolytic enzymes, are easy to isolate, and the existence of the protein in an inactive form allows expression of proteins that are potentially toxic to the host cell (De-Bernardez Clark et al., 1998; Li et al., 2004). Active protein is derived from inclusion bodies by a procedure which starts with the inclusion bodies being released from the cells by mechanical or chemical lysis (Falconer et al., 1999). The inclusion bodies are then isolated by filtration or centrifugation. Purification is achieved via several washing steps, which remove non-specifically bound contaminants, typically cell debris. Purified inclusion bodies are then solubilised using high concentrations of chaotropic reagents such as urea or guanidine hydrochloride (GdHCl), in the presence of reducing agents (i.e. Dithiothreitol (DTT), β-mercaptoethanol (BME)) to break disulfide bonds. After solubilisation, the denatured protein is refolded. This is most commonly achieved by diluting the chaotropic agent, in the presence of redox reagents, to produce native protein. The refolding step often represents a bottleneck and is a step where yields could be improved considerably (Buswell and Middelberg, 2003; De-Bernardez Clark et al., 1998).

Lysozyme has often been used as a model protein to study the effects of operating variables on refolding, as the folding pathway for this protein has been previously characterised in significant detail (Buswell and Middelberg, 2003; Kiefhaber,

1995; Radford et al., 1992; Wildegger and Kiefhaber, 1997). The majority of refolding (approximately 86%) occurs via a single intermediate, whilst a small proportion (approximately 14%) refolds by a fast pathway (Kiefhaber, 1995). The monomolecular refolding reaction competes for the intermediate with a multimolecular aggregation reaction that produces insoluble, inactive aggregates. Previous studies on lysozyme have examined the independent effects of process factors on refolding yield, including GdHCl concentration, redox ratio, protein concentration, mixing intensity, reactor type, pH and inclusion body contaminants (De-Bernardez Clark et al., 1998; Hevehan and De-Bernardez Clark, 1997; Katoh et al., 1999; Katoh and Katoh, 2000; Lee et al., 2002; Maachupalli-Reddy et al., 1997; Mannall et al., 2006; Willis et al., 2005).

The effects of individual parameters on refolding yields have also been characterised for a number of proteins including ribonuclease, trypsinogen, and Fab fragments (Anfinsen and Haber, 1961; Buchner and Rudolph, 1991; Buswell et al., 2002; Fischer et al., 1993). However, few studies have looked at these parameters taken together in a factorial fashion (Buswell et al., 2002). Such a factorial examination can be used to determine the influence of each parameter upon a defined response such as refolding yield, and to quantify the degree of interactions between them. Interactions suggest either positive or negative cooperative effects, beyond the additivity of the parameters alone, and an understanding of them is key to designing optimised processes.

The formation of native protein is favoured at low protein concentrations where the rate of aggregation is minimised (Buswell et al., 2002). Consequently, the

ratio (R); and guanidine hydrochloride concentration (G). The effect of protein concentration was omitted as the decrease in refolding yields at increased protein concentration has been well documented (De-Bernardez Clark et al., 1998; Goldberg et al., 1991; Yasuda et al., 1998). The factorial design experiment varies each component of the refolding solution (guanidine hydrochloride concentration and redox ratio, to change the environment experienced by the denatured protein), and also each of the other factors, both independently and in all possible combinations. Our previous study (Mannall et al., 2006) used a twin impeller to examine the importance of dispersive mixing. The mini-impeller used may provide sufficient dispersion of protein molecules to prevent aggregation, increasing the selectivity of the refolding reaction over that of aggregation. For this reason both impellers were tested for their effect on the refolding reaction. It was not the aim of this chapter to optimise the refolding of lysozyme *per se*, as this has been achieved elsewhere; but rather to understand the relative importance and interactions of the five factors, in fed-batch dilution using the redox system described. An interaction suggests that the effect of a given factor is dependent on the level of another factor. Knowledge of such interactions is critical for the effective design of efficient refolding strategies. It has been shown previously that the redox ratio is critical to refolding yields, with an optimum existing at about 2:1 (reduced : oxidised) for lysozyme (De-Bernardez Clark et al., 1998; Hevehan and De-Bernardez Clark, 1997). This study seeks to establish whether the effect of redox ratio is still significant in a system where the effective redox ratio is defined by the volume of denatured protein added. It also investigates whether the injection rate still remains influential in the presence of very low initial redox ratios.

Observations from the factorial study suggested that it would be valuable to study batch refolding as a comparison to fed batch systems. The second part of this study therefore uses a batch high throughput study to understand how the behaviour of the system changes under batch conditions, and to observe if optimal conditions might be achieved in a simpler, less expensive system. A graphical windows of operation approach is used to determine the influence of conditions on yield and time to yield.

3.2 Materials and Methods

3.2.1 Materials

Chicken egg white lysozyme, lyophilised powder approx. 50,000 units/mg protein, was purchased from Sigma-Aldrich Company Ltd (Poole, Dorset, England, UK).

All additional chemicals were sourced from Sigma-Aldrich Company Ltd (Poole, Dorset, England, UK) and were of at least reagent grade.

3.2.2 Twin Impeller Reactor

Initial experiments used a twin impeller reactor as described previously in Chapter 2 Figure 2.2.

3.2.3 Denaturation-Reduction of Lysozyme

Lysozyme (16 ± 0.1 mg mL^{-1}) was denatured and reduced in 8 M GdHCl, 32 mM DTT, 50 mM Tris, 1 mM EDTA pH 8 for 2 h at 37 °C. Aliquots of denatured lysozyme were stored at –80 °C prior to use. Once thawed, aliquots were not

refrozen. Denatured lysozyme concentrations were measured at 280 nm after 1:100 dilution in acetic acid (0.1 M) and using an extinction coefficient of 2.37 mL cm^{-1} mg^{-1} (lysozyme) (De-Bernardez Clark et al., 1998). The denatured state was confirmed by reverse phase high pressure (performance) liquid chromatography (RP-HPLC).

3.2.4 Reverse Phase HPLC (RP-HPLC)

Assessment of native and non-native lysozyme concentration was achieved using the RP-HPLC method described earlier in section 2.2.11 based upon that of Lee et al (Lee et al., 2002).. Native lysozyme had a retention time of 8.57 ± 0.03 min, denatured lysozyme had a retention time of 11.20 ± 0.1 min.

3.2.5 Effect of Guanidine Hydrochloride Concentration

Denatured-reduced lysozyme (16 mg mL^{-1}) was diluted 1:15 in various refold buffers (six refold buffers in total as detailed in Table 3.1) at ambient temperature and mixed vigorously for 30 seconds, and left at ambient temperature. After 2 h a sample (2 mL) was taken and quenched with 10% v/v TFA (200 µL). A further sample was taken at 24 h and left unquenched. Samples were assayed for native lysozyme using RP-HPLC.

Buffer	[Tris.HCl] (mM)	[Cystamine] (mM)	[EDTA] (mM)	[GdHCl] (M)	Final [GdHCl] (M)	pH
1	50	4	1	0	0.53	8.0
2	50	4	1	0.167	0.69	8.0
3	50	4	1	0.47	0.97	8.0
4	50	4	1	0.717	1.20	8.0
5	50	4	1	0.967	1.44	8.0
6	50	4	1	1.967	2.38	8.0

Table 3.1: Compositions of buffers utilised when studying the effect of GdHCl concentration on refolding yields.

3.2.6 Factorial Experiments

Denatured-reduced lysozyme (16 mg mL^{-1}) was refolded by 1:15 dilution at 25 °C, once for each of the combinations of different mini-impeller intensities (Re_m), bulk- impeller intensities (Re_b), injection rates (Q), redox ratios (R) and final GdHCl concentrations (G) as detailed in Table 3.3. Eight midpoints were performed, to assess the levels of variance in the generated data. In all cases samples were mixed for a total of 2 h, and then left for a further 22 h at ambient temperature. 2 mL samples were removed and quenched at 2 h and 24 h with 200 μL of 10 % (v/v) TFA. Samples were assayed by RP-HPLC to determine refolding yields. Five refold buffers were prepared as detailed in Table 3.2. Factor levels used in the factorial experiments are detailed in Table 3.3. Analysis of the factorial data was carried out so as to calculate the effect of each factor at a 95 % confidence interval. Confidence intervals were calculated according section 3.2.8.

An estimate of individual effects of factor was calculated as the difference between the mean yield of all the runs when the factor was present, subtracted from the mean of the yield of all runs when the factor was absent. 2 factor interactions were calculated as the difference between the mean yield for all runs where both factors present are at a high or low level (++,--) and the mean of all runs where only one of the factors is present (+-,-+).

Buffer	[Tris.HCl] (mM)	[Cystamine] (mM)	Redox ratio	[EDTA] (mM)	[GdHCl] (M)	Final [GdHCl] (M)	pH
1	50	4	2:1 (+)	1	0.717	1.2 (+)	8.0
2	50	4	2:1 (+)	1	0	0.533 (-)	8.0
3	50	25	0.17:1 (-)	1	0.717	1.2 (+)	8.0
4	50	25	0.17:1 (-)	1	0	0.533 (-)	8.0
5	50	5.67	1.085:1 (0)	1	0.358	0.867 (0)	8.0

Table 3.2: Compositions of buffers used for factorial experiments. (+) indicates upper level of factor, (-) indicates lower level of factor, (0) indicates the midpoint.

Factor	Upper (+)	Lower (-)	Midpoint (0)
Re_b	576	3836	2206
Re_m	2000	0	1000
Q	1.23 mL min^{-1}	0.115 mL min^{-1}	0.669 mL min^{-1}
R	2:1	0.17:1	1.085:1
G	1.2 M	0.533 M	0.866 M

Table 3.3: Levels used in factorial experiments. Values were set on the basis of initial screening experiments. Re_b bulk impeller Reynolds number, Re_m mini-impeller Reynolds number, Q injection rate (mL min^{-1}), R Redox ratio, G Final GdHCl concentration (M).

3.2.7 Time to Yield Effects for Factorial Experiment

Yields from 2 h and 24 h were used to estimate approximate time required to reach the maximal refolding yield, assumed to be that achieved at 24 hours. Exponential fits (3.1) to the refolding data were used to estimate approximate refolding time.

$$y = a(1 - e^{bt}) \tag{3.1}$$

where a and b are constants, y is yield and t is refold time in hours.

Effects of factors were calculated as described in section 3.2.6, confidence intervals were calculated according section 3.2.8.

3.2.8 Calculation of Confidence Intervals for Factorial Experiments

The residual standard deviation (RSD) is used as an estimate of the standard deviation of the data. This is subsequently used to calculate confidence intervals. RSD is calculated using equation (3.2):

$$RSD = \sqrt{\frac{\Sigma(Y_{obs} - Y_{predicted})^2}{n - p}} \qquad (3.2)$$

Where: Y_{obs}= Observed value under a particular condition, $Y_{predicted}$= value predicted from formula (3.3) under a particular condition, n= total number of experiments, p= number of terms in the model (3.3).

Predicted values are calculated from a linear model fitted to the data, described by formula (3.3):

$$Y = \beta_0 + \beta_{Re_b} x_{Re_b} + \beta_{Re_m} x_{Re_m} + \beta_Q x_Q + \beta_R x_R + \beta_G x_G + \beta_{Re_b Re_m} x_{Re_b} x_{Re_m} + \beta_{Re_b Q} x_{Re_b} x_Q$$

$$+ \beta_{Re_b R} x_{Re_b} x_R + \beta_{Re_b G} x_{Re_b} x_G + \beta_{Re_m Q} x_{Re_m} x_Q + \beta_{Re_m R} x_{Re_m} x_R + \beta_{Re_m G} x_{Re_m} x_G + \beta_{QR} x_Q x_R$$

$$+ \beta_{QG} x_Q x_G + \beta_{RG} x_R x_G \qquad (3.3)$$

Where: β_x= parameter vector for factor x, x_A=variable for factor A, which takes a value of +1 for an upper level of a factor, -1 for a lower levels of factor and 0 for a midpoint. The model is used for screening experiments to gain an appreciation of important factors and interactions, for this reason it is limited to two factor interactions. Calculation of vector parameters is achieved by multiple regression using the design matrix (X) and the vector of observations (Y). Each of the design matrix rows values corresponds to the variable value for each parameter

vector under a particular combination of conditions. This combination of conditions produces an observation, which is given in Y. Estimate of parameter vectors is achieved by a least square method according to formula (3.4):

$$\hat{\beta} = (X^T X)^{-1} X^T Y \qquad (3.4)$$

Where X= design matrix, X^T= design matrix transformed, Y= vector of observations.

These estimated values are then placed in to the model and values of $Y_{predicted}$ calculated for each of the combinations of factor used. These predicted values are then used in combination with observed values to calculate the RSD according to formula (3.2). The degrees of freedom for RSD is defined by (n-p), there are 16 terms in the model described above, hence this experiment has (40-16) 24 degrees of freedom.

The confidence interval is calculated from the RSD. RSD is converted to s_{eff} according to formula (3.5):

$$s_{eff} = \frac{RSD}{\sqrt{2^{k-2} r}} \qquad (3.5)$$

Where r is the number of replicates of the design. This is converted to confidence interval for the effects by formula (3.6):

$$A = t_{\nu} s_{eff} \qquad (3.6)$$

Where t_ν= value of t at P=0.05 under the degrees of freedom ν.

It must be noted that the residual standard deviation incorporates both lack of fit of the model (calculated from residuals in the 2k design) and the pure error from the midpoint replicates. The sum of squares of the residuals is equal to sum of the sum of squares of the lack of fit (v=17) and those of the midpoint replicates.

Significant factors described in this chapter are significant to a 95% confidence interval calculated with RSD.

3.2.9 Batch Refolding Tests For Yield and Throughput at Various Redox-Ratio and GdHCl Concentrations

The fed-batch factorial study suggested that high refolding yields could be achieved even with rapid injection and poor mixing efficiency. Comparative studies of the influence of two key parameters, redox ratio and GdHCl concentration, were therefore carried out for a simple batch refolding system. In this system, the final yields obtained for each condition, and the time taken to reach these yields, were both determined. Each refold was left for a total of 24 h to allow the reaction to proceed to practical completion. For the purposes of subsequent analysis, the maximum practicable yield achieved under each condition was assumed to be equivalent to that measured at 24 h.

Denatured and reduced lysozyme (15.9 mg mL^{-1}) was refolded by 1:15 dilution into a series of buffers of varying GdHCl concentrations and redox ratios. The denatured protein was added in a single step and distributed by a magnetic stirrer for 5 minutes to ensure homogeneity. Samples (2 mL) were taken at 1, 2, 4, 8 and 24 h and quenched with 200 µL of 10 % (v/v) TFA. Nine buffers were prepared as detailed in Table 3.4.

Buffer	[Tris.HCl] (mM)	[Cystamine] (mM)	Redox ratio	[EDTA] (mM)	[GdHCl] (M)	Final [GdHCl] (M)	pH
1	50	4	2:1	1	0	0.533	8.0
2	50	4	2:1	1	0.358	0.867	8.0
3	50	4	2:1	1	0.717	1.2	8.0
4	50	5.67	1.085:1	1	0	0.533	8.0
5	50	5.67	1.085:1	1	0.358	0.867	8.0
6	50	5.67	1.085:1	1	0.717	1.2	8.0
7	50	25	0.17:1	1	0	0.533	8.0
8	50	25	0.17:1	1	0.358	0.867	8.0
9	50	25	0.17:1	1	0.717	1.2	8.0

Table 3.4: Compositions of buffers used for windows of operation experiments

The maximum yield and the time to achieve maximum yield were used to produce contour plots. Exponential fits to the data using equation (3.1), were used to predict the time taken to achieve specific yields.

Contour plots were generated from the data for both yield and time to yield. The experimental data generated was fitted using the following expression so as to account for the possible effects and interaction between redox ratio and guanidine hydrochloride concentration (Equation 3.9).

$$z = b_1 + b_2 R + b_3 G + b_4 R^2 + b_5 G^2 + b_6 RG \qquad (3.9)$$

where z represents yield or time to yield, R is redox ratio, G is guanidine concentration. The constants b_1-b_6 were determined by multiple regression.

The resulting contour plots predict the effect of combinations of operating conditions. They were used in this study to generate windows of operation to examine the trade-offs between the required levels of yield and the associated process time required to achieve these.

3.3 Results and Discussion

3.3.1 Effect of Guanidine Hydrochloride Concentration on Refolding Yields

Of the five factors used in the factorial study, appropriate upper and lower levels of final GdHCl concentration were determined by experiment. Final chaotrope concentration in the reactor has been shown to affect yields in the refolding of lysozyme (De-Bernardez Clark et al., 1998; Hevehan and De-Bernardez Clark, 1997; Katoh et al., 1999; Lee et al., 2002). It is believed that high concentrations of GdHCl decrease both the rate of refolding and aggregation, but that aggregation is the more severely affected (De-Bernardez Clark et al., 1998). Therefore, denaturant concentration must be low enough to permit protein molecules to refold, but high enough to prevent protein aggregation and to promote protein flexibility which is required for structural reorganisation (Tsumoto et al., 2003). An optimal concentration therefore exists where yields are maximised without significantly reducing the rate of refolding (De-Bernardez Clark et al., 1998).

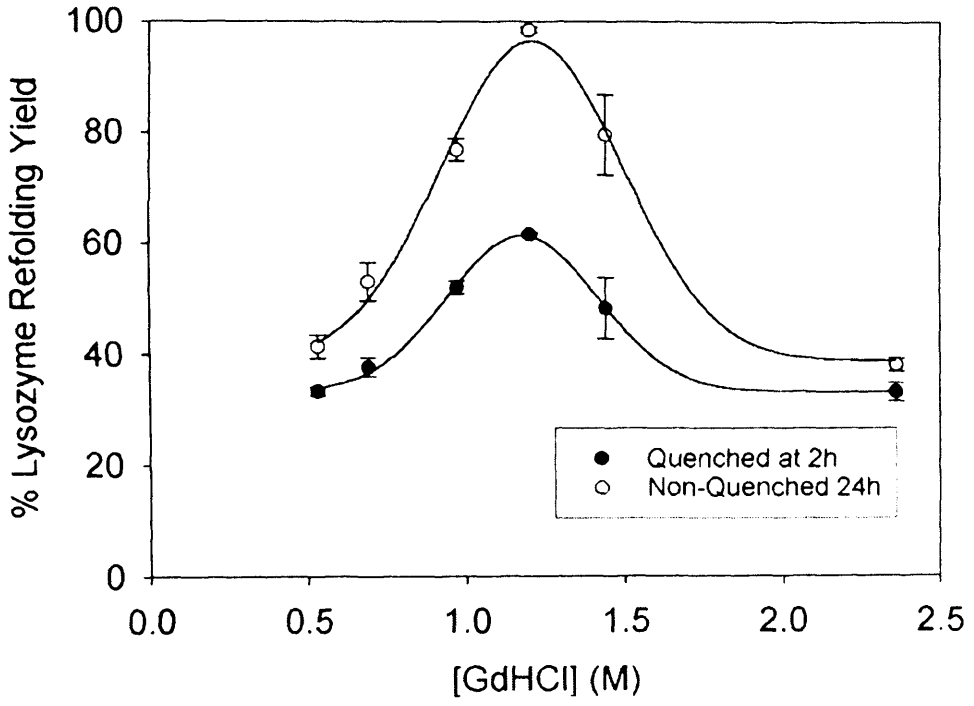


Figure 3.1: Effect of GdHCl concentration upon refolding yield of lysozyme at a final concentration of 1.1 mg mL^{-1}, quenched at 2 h (•) and 24 h (○). The optimal GdHCl concentration appears to be 1.2 M with any further increase in GdHCl concentration reducing refolding yields.

In setting the limits for final GdHCl concentration in the refold buffer, it is evident that an upper limit of approximately 4.2 M exists. This is defined by the $C_{0.5}$ of lysozyme in GdHCl (the concentration of chaotrope at which half the protein is folded) (Ahmad and Bigelow, 1982). This was measured at pH 7 rather than pH 8 (used in this study) and is therefore only an approximation of the $C_{0.5}$ under conditions used in these experiments The yields of refolded lysozyme at various final GdHCl concentrations can be seen in Figure 3.1, at 2 h (quenched) and 24 h (unquenched) after the start of the refold step. The optimal final GdHCl concentration is 1.2 ± 0.1 M, which is in close agreement with the 1.3 M optimum found by others (De-Bernardez Clark et al., 1998). Although we used the same

levels of Tris and pH, the slight difference might be accounted for by the difference in redox ratios used (2:1 here versus approximately 1.33:1 in previous work) and the use of cystamine in place of oxidised glutahthione. As discussed by these authors, the increasing yield up to 1.2 M GdHCl is predominantly due to a decreasing rate of aggregation, while the lower yields achieved above 1.2 M GdHCl concentrations are more likely due to the considerably slowed refolding rates. Extrapolation of refolding kinetic data from our labs suggests it may take 3 days for the reaction to go to completion at a final GdHCl concentration of 2.36 M (data not shown). In reactions where the GdHCl concentration was above the optimal 1.2 M, no aggregation was visible, although a series of peaks prior to the native peak were found on RP-HPLC analysis as observed in Figure 3.2. The combined areas of these peaks are presented in Figure 3.3 and most likely represent soluble misfolded protein, formed by incorrect disulfide pairings and may include mixed disulfides though the possibility of these peaks representing soluble aggregates cannot be discounted as the methods used here cannot identify the true identity of these species. Activity measurements would likewise not be able to identify these species. Comparison of RP-HPLC and activity measurement (Buswell and Middelberg, 2002) suggest that activity is associated with a non native peak. The amount of these species appeared to increase with GdHCl concentration, with a sharp increase at 1.44 M GdHCl. Refolding of protein above 1.2 M GdHCl may be so slow that these incorrectly folded forms can accumulate.

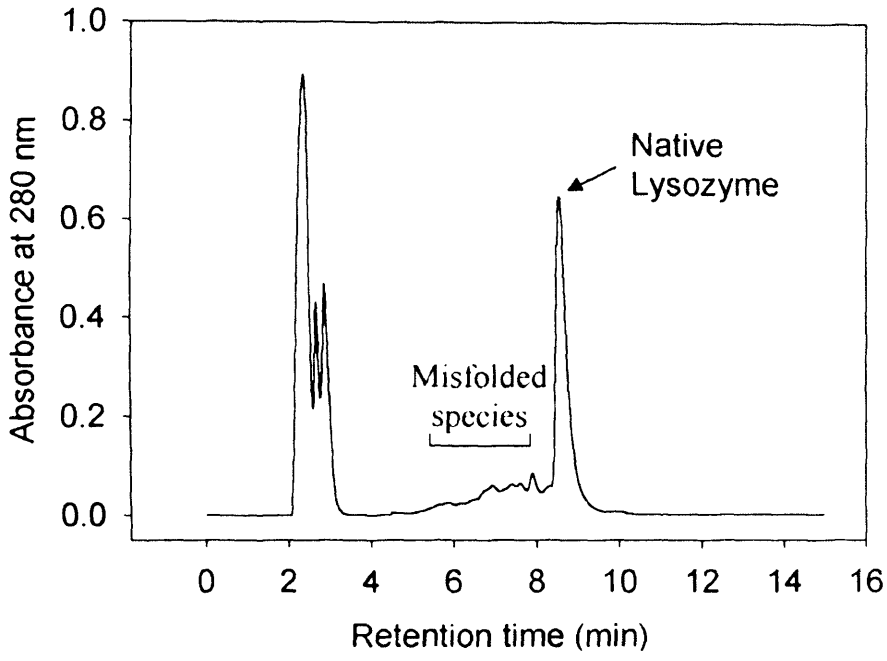


Figure 3.2: RP-HPLC trace, showing the retention of native protein (8.57 mins) and misfolded species (≈ 6-8 mins), presumed to be disulfide variants of the native structure.

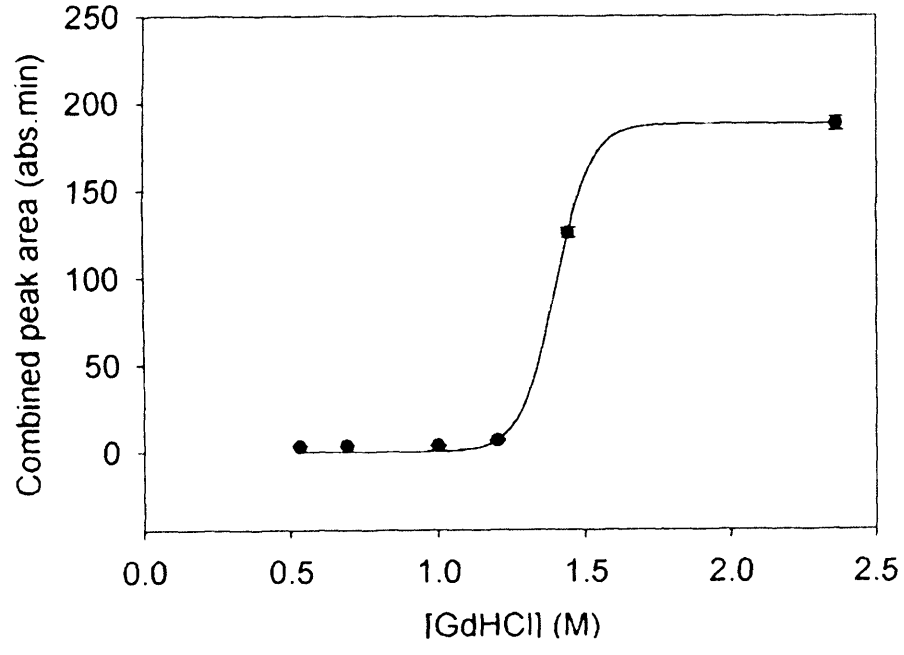


Figure 3.3: Effect of GdHCl concentration on the amount of misfolded protein species from the refolding reaction quenched at 2 h. The amount increases as the concentration of GdHCl increases with a sharp increase at 1.44 M.

The unconventional redox system used in this study may also have an impact on the formation of incorrectly-folded proteins. The refolding buffer initially contains only the oxidised form of the redox reagent (cystamine), whereas the denatured protein contains only an excess of reducing agent (DTT). High levels of the oxidising species present during the initial stages of injection may cause rapid oxidation of free thiols, resulting in a greater number of incorrect disulfide-bond pairings than when using a conventional refolding buffer containing initially both redox components. The magnitude of this effect is expected to be inversely proportional to injection rate with the formation of misformed disulfides favoured at slower injection rates.

3.3.2 Factorial Experiments

Results from the factorial experiments are detailed in Table 3.5 and were used to calculate the impact of the factors upon refolding yields, which are summarised in Figures 3.4 and 3.5. Discussion of these effects first considers significant single factor effects and then the most significant interactions.

3.3.2.1 Effects and Interactions

Final GdHCl Concentration

Figures 3.4 and 3.5 show that final GdHCl concentration (G) influences the refolding yield measured at both 2 h and 24 h, in a similar fashion to that noted in previous studies (Katoh et al., 1999), where increases in chaotrope concentration raised the refolding yield. High concentrations (1.2 M) of GdHCl will prevent aggregation of the protein, leading to greater amounts of soluble protein being present, which under appropriate redox conditions will be native. The effect of GdHCl concentration upon refolding yield appears to be significantly reduced in

the presence of a low redox ratio (0.17:1) (Table 3.5). This is confirmed by the presence of a significant interaction between these two factors (R*G) shown in Figures 3.4 and 3.5. High levels of GdHCl (1.2 M) keep the protein from aggregating, but operating at a low redox ratio (0.17:1) will result in an increase of mismatched/mixed disulfide species being present and a limited ability to shuffle these to the native form, leading to reduced final yields. These assertions appear to be confirmed by RP-HPLC analysis of high GdHCl concentration at low redox refolds, where a series of peaks prior to the native peak are clearly visible (Figure 3.2), and increase in size at higher GdHCl concentrations (Figure 3.3). These peaks may represent the soluble mis-paired disulfide bonded species, including mixed disulfides as assumed in a previous study (De-Bernardez Clark et al., 1998) where at low redox ratios, low yields could not be explained by aggregation alone but were presumed to be caused by glutathionation of folding intermediates. In the current study the folding intermediates may have been cysteaminated. As discussed above, the possibility of these peaks representing soluble aggregates cannot be discounted, as the methods used here cannot identify the true identity of these species.

Redox Ratio

As with previous work (De-Bernardez Clark et al., 1998; Hevehan and De-Bernardez Clark, 1997) it appears that the redox ratio (R) of reduced to oxidised species of the redox reagent, has a significant effect on refolding yield at both 2 and 24 h (Figures 3.4 & 3.5), even though the redox ratio changes considerably over the injection period. The two redox species act as a redox pair, allowing the shuffling of disulfides until the correct pairings are achieved. An optimal ratio is

required to allow efficient disulfide shuffling. At low redox ratios (0.17:1) the concentration of cystamine (oxidised species) is too high, particularly in the initial injection period, to allow efficient disulfide exchange. This may lead to the rapid formation and trapping of incorrectly disulfide-paired proteins and reduced final yields. The effect of redox ratio is most obvious at high GdHCl concentrations (1.2 M), and an interaction between these two factors (R*G) is confirmed in Figures 3.4 and 3.5. The lower rate of aggregation at high GdHCl affords the redox ratio present more time to promote the formation of correctly paired cysteines and hence the formation of native protein. By contrast, the rapid aggregation of protein at the low GdHCl concentration (0.533 M) decreases the time for which the effect of redox ratio can act and hence also decreases its impact upon refolding yields.

Rate of Injection

It has been shown previously that refolding yields can be improved by decreasing the rate of injection of the denatured protein (Katoh et al., 1999; Katoh and Katoh, 2000). The formation of native protein is favoured at low protein concentrations (Buswell et al., 2002) where aggregation of folding intermediates is slower than refolding to the native protein. Similarly, the gradual addition of denatured protein to refolding buffer results in improved yields due to the presence of a lower concentration of partially refolded protein over the duration of the addition (Katoh et al ., 1999).

From Figures 3.4 and 3.5 it is observed that increased injection rates (Q) have a significant negative effect upon the refolding yields measured at 24 h, but has a

positive effect at 2 h. There is a considerable difference between the injection conditions used. Under a slow injection rate (0.115 mL min^{-1}) complete addition takes just under 2 hours (≈ 1 h 55 min), whilst under a fast injection rate (1.23 mL min^{-1}) complete addition takes just over ten minutes. This means that under fast injection conditions, the protein is at the final protein concentration for about an hour and fifty minutes before the 2h sample is taken, whereas under slow injection conditions it is at this concentration for only 5 minutes. Therefore, for the two-hour time point, the fast injection rate is likely to give a greater yield than the slow injection rate, given the greater time available to finish refolding, after complete addition of the protein. However, increased injection rates have a significant negative effect upon yield at 24 h. Slow injection of the denatured protein minimises the effective concentration of partially renatured intermediates that are present at any specific time. Therefore, the denatured protein entering the system is less likely to collide to form aggregates. It is surprising that an effect of injection rate is found in such a system where the protein initially experiences very low redox ratios upon injection. The apparent solubility of some of the mis-disulfide bonded species that are more likely to be formed under high redox ratio and an ability to reshuffle these forms, means that protein is not rapidly lost by this route. The insignificant loss of protein here may explain why low injection rates still give greater yields despite the very low redox ratios experienced. The effect of injection rate appears to become less significant at higher GdHCl concentrations (Table 3.5); an effect confirmed by a significant interaction between these two factors (Q*G) shown in Figure 3.5. As observed in Figure 3.1, little aggregation is observed, under batch conditions, at high GdHCl concentrations (≥1.2 M). In the fed-batch system employed in the study, GdHCl

concentration and protein concentration are directly related. Over the addition period, GdHCl may be present at concentrations sufficient to slow the rate of aggregation, preventing precipitation at these protein concentrations. This effect appears to apply at all injection rates and possibly explains the diminished effect of injection rate upon refolding yields at high GdHCl concentrations.

Cond no.	Re_b	Re_m	Q	R	G	Y@2 h	Y@24 h
1	+	+	+	+	+	78	92
2	+	+	+	+	-	41	43
3	+	+	+	-	+	36	52
4	+	+	+	-	-	30	34
5	+	+	-	+	+	46	96
6	+	+	-	+	-	46	69
7	+	+	-	-	+	27	50
8	+	+	-	-	-	33	55
9	+	-	+	+	+	72	93
10	+	-	+	+	-	38	41
11	+	-	+	-	+	37	53
12	+	-	+	-	-	33	40
13	+	-	-	+	+	51	100
14	+	-	-	+	-	43	60
15	+	-	-	-	+	29	48
16	+	-	-	-	-	34	57
17	-	+	+	+	+	76	92
18	-	+	+	+	-	36	39
19	-	+	+	-	+	35	50
20	-	+	+	-	-	32	41
21	-	+	-	+	+	54	100
22	-	+	-	+	-	43	65
23	-	+	-	-	+	32	67
24	-	+	-	-	-	28	49
25	-	-	+	+	+	70	94
26	-	-	+	+	-	38	40
27	-	-	+	-	+	41	63
28	-	-	+	-	-	33	42
29	-	-	-	+	+	48	100
30	-	-	-	+	-	42	63
31	-	-	-	-	+	33	70
32	-	-	-	-	-	30	51
33	0	0	0	0	0	56	70
34	0	0	0	0	0	50	75
35	0	0	0	0	0	57	69
36	0	0	0	0	0	58	67
37	0	0	0	0	0	59	73
38	0	0	0	0	0	62	77
39	0	0	0	0	0	62	70
40	0	0	0	0	0	60	73

Table 3.5: 2^5 factorial experimental design. Each combination was repeated once, and 8 midpoints were performed to provide an estimate of the variance in the data.

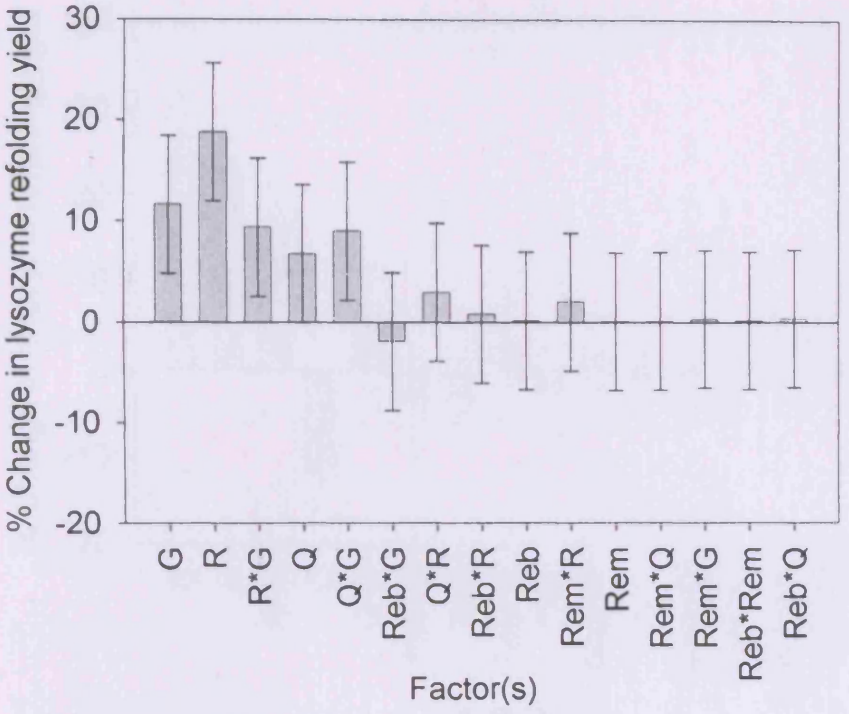


Figure 3.4: Effect of factors for reaction samples quenched at 2 h, results are shown for a 95 % confidence interval. Significant single factors GdHCl concentration G and redox ratio R. Significant two factor interactions were found between GdHCl concentration G and redox ratio R, and between injection rate Q and guanidine concentration G. None of the factors/interactions appear to be more significant than others. Error bars indicate 95 % confidence interval.

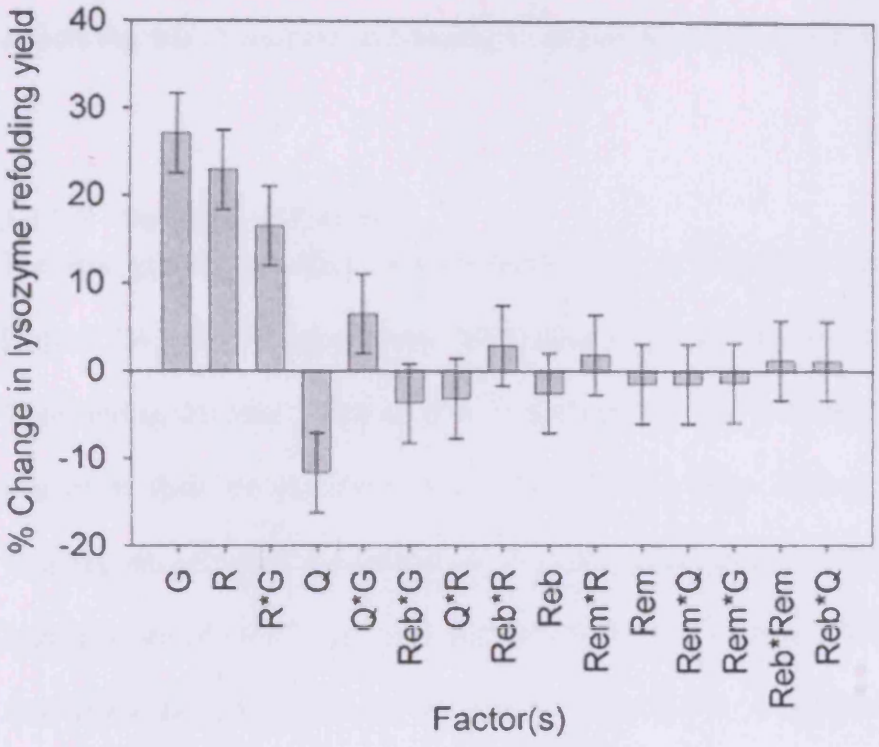


Figure 3.5: Effect of factors for reaction samples quenched at 24 h, results are shown for a 95 % confidence interval. Significant single factors GdHCl concentration G and redox ratio R. Significant two factor interactions were found between GdHCl concentration G and redox ratio R, and between injection rate Q and guanidine concentration G. None of the factors/interactions appear to be more significant than others. Error bars indicate 95 % confidence interval.

3.3.2.2 Absence of Positive Effect of Mixing

It has been shown previously that mixing is unimportant at higher chaotrope concentrations (Lee et al., 2002). It appears that under the GdHCl concentrations used in this fed-batch study, mixing had only a limited effect upon protein refolding yield. This may be explained by the reduced aggregation rate at higher GdHCl concentrations. For mixing to affect the rate of reaction it must operate at a slower or a similar rate to that of the reaction. At high GdHCl concentrations the rate of aggregation and refolding reactions may be reduced to levels

considerably below that of the mixing rate. If this is the case mixing no longer affects the rate of reaction and hence the degree of refolding achieved.

3.3.2.3 Strength of Effects

The strength of the effect of each factor may be judged at 2 h and 24 h from Figures 3.4 and 3.5, respectively. For yields at 2 h no effect appears to dominate. Considering the final yields at 24 h, it appears that GdHCl concentration is more important than the injection rate. The injection rate reduces aggregation by limiting the effective concentration of aggregating species in the reactor, whilst higher levels of GdHCl prevent aggregation by maintaining solubility. It appears that under the limits used in this factorial experiment, prevention of aggregation by manipulating the chemical environment has a greater impact than relying on physical means afforded by a reduced injection rate.

3.3.3 Time to Yield for Factorial Experiment

It is clear that from Figure 3.6 that injection rate has a significant negative effect upon time to yield. It is evident that the longer it takes to add the denatured protein, the longer it would take to refold. Under slow injection times the final protein concentration is reached more slowly and hence comparatively the refolding protein has less time in which to refold.

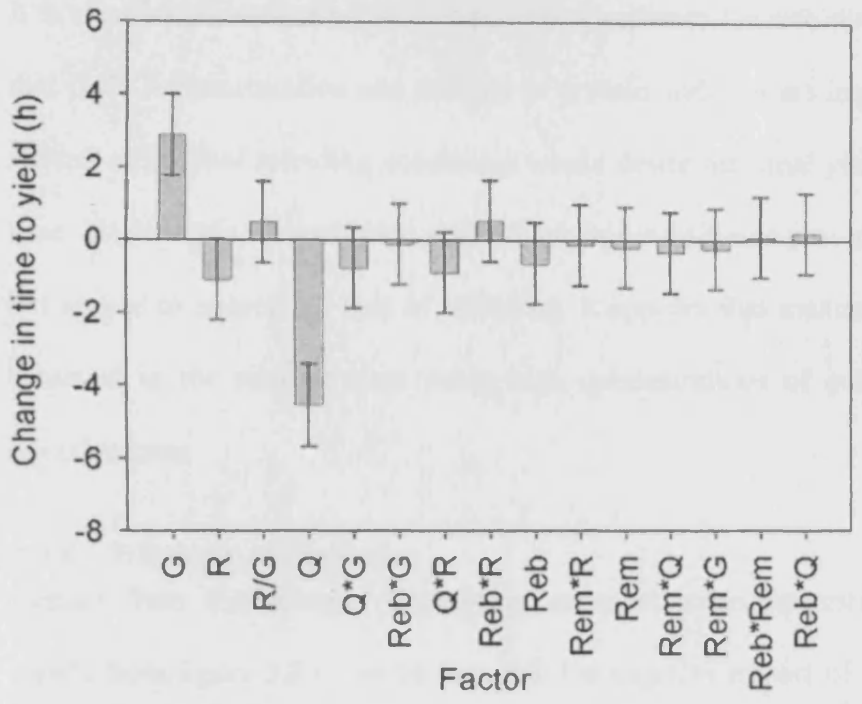


Figure 3.6: Effects plot for time to yield. It is clear that there is a significant positive effect of GdHCl concentration and a significant negative effect of denatured protein flow rate. Error bars indicate 95% confidence interval.

Another critical factor from Figure 3.6 appears to be GdHCl concentration. It appears that GdHCl concentration has a significant effect upon time to yield. It appears that those runs where GdHCl concentrations are high take longer to refold. On average those refolds conducted with high GdHCl concentrations give greater yields. Larger amounts may take longer to refold. Under high levels of GdHCl protein is less likely to aggregate, and will have longer to refold.

Redox ratio appears to have a marginal negative effect on refolding time. This effect is particularly apparent at high GdHCl concentrations where the redox ratio has time to take effect. Lower redox ratios will take longer to refold.

It is clear when comparing factors that are significant for refolding yield and rate that GdHCl concentration and the rate of protein addition are important for both. Pursuit of optimal refolding conditions would desire maximal yields in a minimal time. High levels of guanidine and slower injection times provide higher yields, but appear to reduce the rate of refolding. It appears that maximal yield may be achieved in the shortest time using high concentrations of guanidine and fast injection rates.

3.3.4 Windows of Operation

Results from the factorial experiment revealed some interesting interactions. Firstly from figure 3.5 it can be seen that the negative impact of rapidly injecting the denatured protein is diminished at high GdHCl concentrations (at and above 1.2 M). This suggests that high yields may be generated even at fast injection rates such as those achieved in a simple batch mode of operation. Under certain conditions overall yields appeared to decrease with increased mixing efficiency. In addition it was found that injection rate has a significant negative effect on time to yield. Taken together these results suggest that behaviour of the system could well be different under batch conditions where addition is very rapid and a minimum amount of mixing is employed. A batch refolding experiment was therefore set up to study the effect of guanidine concentration and redox ratio, the two most important factors, under batch addition conditions with limited mixing.

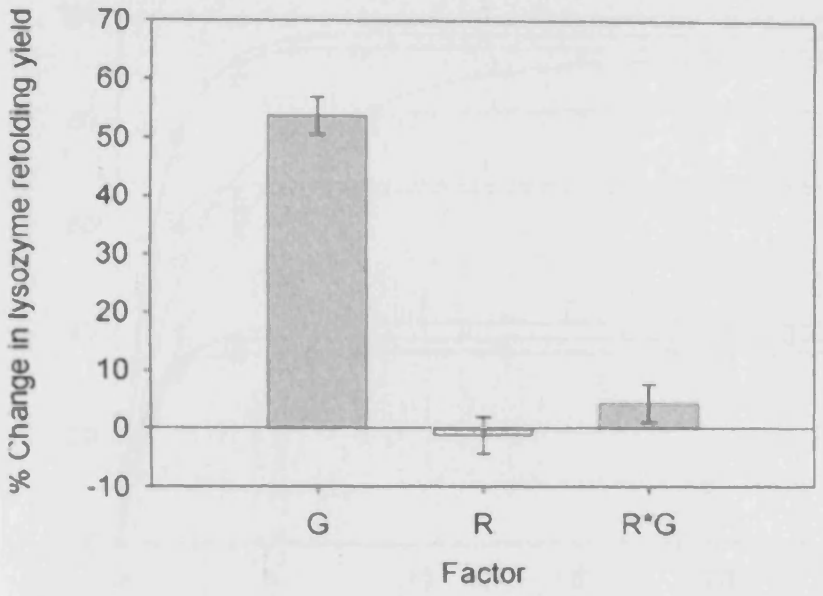


Figure 3.7: Effects plot for batch refolding. It appears that the effect of redox ratio has decreased considerably in comparison to fed batch.

Initial results shown in the effects plot (Figure 3.7) suggested that by increasing the injection rate and limiting mixing, the effect of redox ratio at 24 h was lessened considerably. In particular, the effect of redox ratio was considerably lower at high guanidine concentrations when compared to the fed-batch system. Refolding curves under various conditions (Figure 3.8) show that with batch refolding, high yields may be achieved but take longer to reach their maxima. Data in Figure 3.8 were used to create contour plots of maximal yield of refolded protein. The times taken to achieve these yields are given in Figures 3.9A and 3.9B. The plots reveal that the greatest yields are achieved with high GdHCl concentrations and that the fastest reactions occur at higher redox ratios. A windows of operation approach (Woodley and Titchener Hooker, 1996) was used to visualise the effects of reducing agent levels and guanidine concentration on refold yields and time to yield.

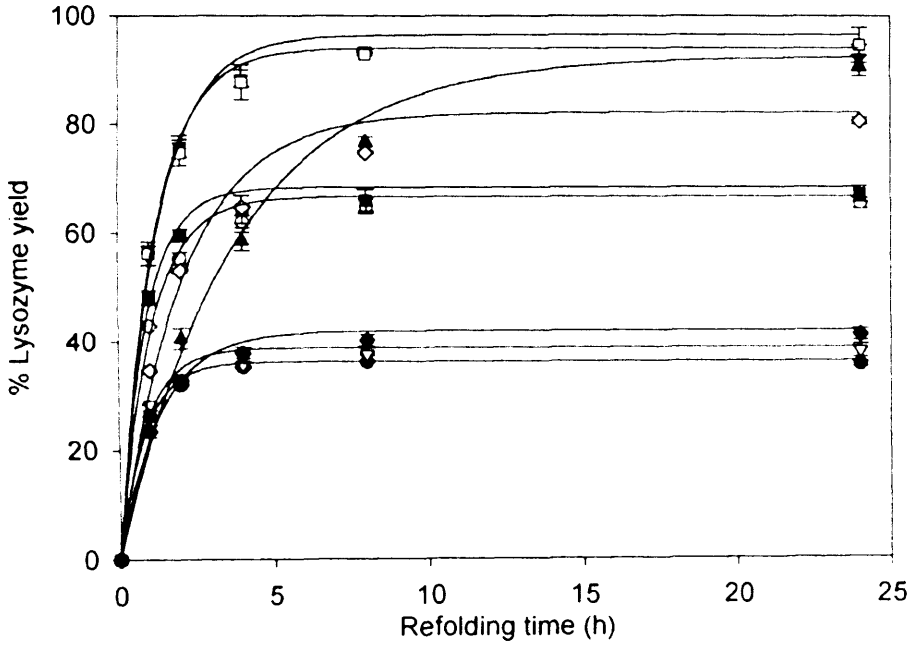


Figure 3.8: Refolding curves under a range of Guanidine hydrochloride concentrations (GdHCl) and redox ratios (RR). (●) 0.533 M GdHCl, 2:1 RR, (○) 0.866 M GdHCl, 2:1 RR, (▼) 1.2 M GdHCl, 2:1 RR, (▽) 0.533 M GdHCl, 1:1 RR, (■) 0.866 M GdHCl, 1:1 RR, (□) 1.2 M GdHCl, 1:1 RR , (♦) 0.533 M GdHCl, 0.17:1 RR, (◊) 0.866 M GdHCl, 0.17:1 RR, (▲) 1.2 M GdHCl, 0.17:1 RR. Curves fitted using exponential relationship (1), where 'a' is the maximum yield (R^2 > 0.987). The greatest yields are achieved at high GdHCl concentrations. At the lowest redox ratio conversion appears to be slow. This is especially marked at 0.866 M GdHCl and above.

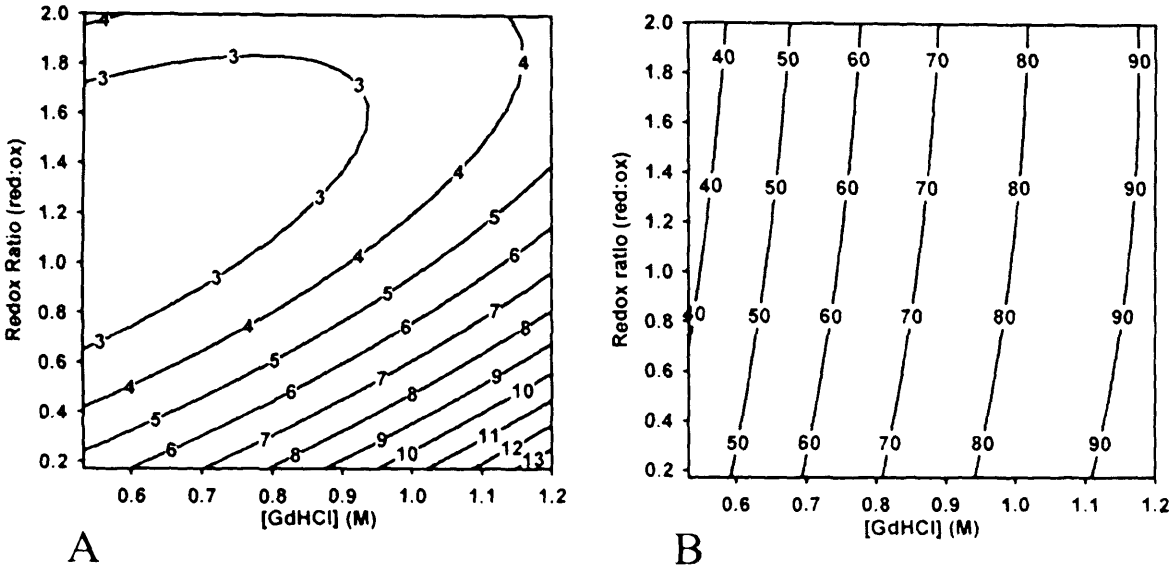


Figure 3.9: (A) Contour plot of time to maximum yield for a variety of redox ratios and guanidine hydrochloride concentrations. Lines link equal time values. It appears that the slowest reactions are achieved at low redox ratio and high guanidine concentrations. The fastest refolding is achieved at redox ratios of one and above. (B) Contour plot of maximum yield as a function of redox ratio and guanidine concentration lines link equal % yield values. It is clear that the GdHCl has the clearest effect on yield in the batch refolding.

From these windows it is possible to determine the combinations of conditions needed to realise specific yields in set time periods. In industry, process time is critical. Therefore, the first simulation was performed to determine conditions that achieve 90 % maximal yield within a typical 8 h shift period. These criteria yield Figure 3.10A, where high guanidine concentrations (> 1.1 M) and redox ratios (≥ 0.7) are indicated. Tightening of the time constraint to 6 h generates Figure 3.10B, where a window exists such that the two criteria are fulfilled at yet higher redox ratios (> 1) and high GdHCl concentrations (> 1.1 M). Tightening the time constraints still further to 4 h or less generates Figure 3.10C. It is

apparent from this plot that the two process specifications cannot be achieved simultaneously. Slackening of the limits on yield to 80 %, while maintaining the time to yield at 4 h, generates Figure 3.10D and a feasible operating window that encompasses lower concentrations of GdHCl. In practice this small saving on GdHCl would decrease costs slightly, but the associated drop in yield would need to be balanced against this before a final decision as to the desirability of operating under these conditions could be made.

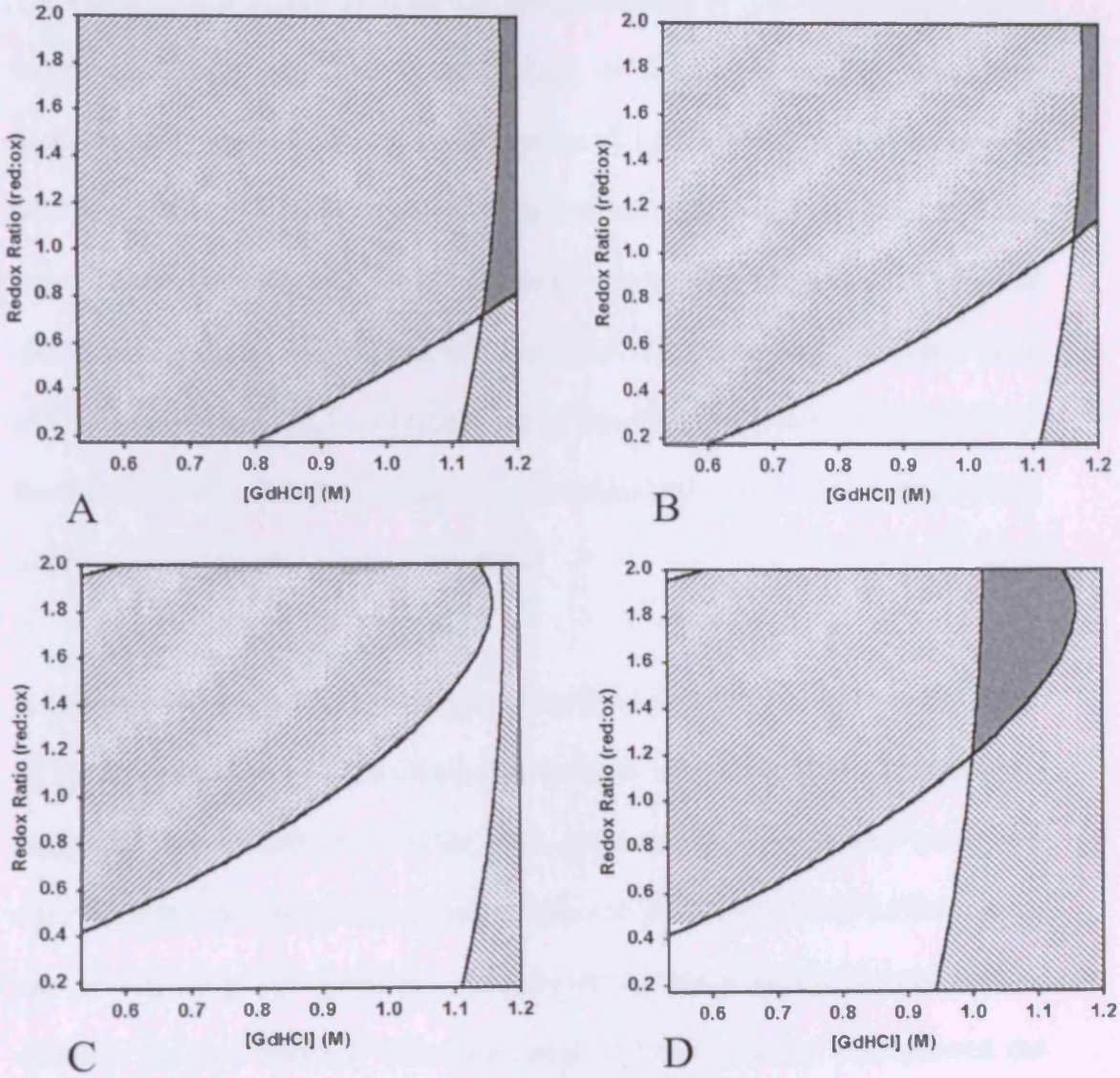


Figure 3.10: Windows of operation for a range of process specifications experienced in terms of yield and process time. Conditions that satisfy yield specifications: ▧. Conditions that satisfy process time constraints: ▨. Windows of operation formed when both conditions are met satisfactorily and simultaneously: ■. (A) Window of operation for ≥ 90 % in 8 h or less. (B) Window of operation for yield ≥ 90 % in 6 h or less. (C) Window of operation for yield ≥ 90 % achieved in 4 h or less. (D) Window of operation for yields ≥ 80 % maximum yield achieved reached in 4 h or less.

3.4 Conclusions

The effect of five factors upon the fed-batch refolding of lysozyme using a refold buffer containing only the oxidised form of the redox reagent have been experimentally studied. A final concentration of 1.2 M GdHCl maximises protein refolding yields. GdHCl concentration and redox ratio have the most significant impact upon the refolding yield. Injection rate is also important. The study identified significant interactions between GdHCl concentration and redox ratio, and between final GdHCl concentration and injection rate. Further analysis of the factorial data revealed that GdHCl concentration and injection-rate were critical to determining the time required to refold.

A comparative study was undertaken using batch refolding to observe the effect of the key parameters of guanidine concentration and redox ratio. This analysis suggested that the effect of redox ratio upon refolding yield was much less significant in this system, though redox ratio still had a considerable effect on the rate of refolding. A graphical windows of operation approach was used to visualise the data obtained from this series of batch refolds. This allowed the appropriate balance between yield and process time to be visualised, and for the conditions of GdHCl concentration and redox ratio that are necessary to realise desired levels of process performance to be identified.

The thesis so far has considered the effect of process factors upon refolding of a pure protein. Inefficiencies in IB isolation mean that refolding typically occurs from impure protein. Chapter 4 therefore analyses the effect of inclusion body purification processes upon eventual refolding yield.

4 The Effect of Contaminants Upon Refolding of Trypsinogen.

4.1 Introduction

Expression of recombinant proteins in bacteria such as *E.coli* often results in the formation of inclusion bodies (IBs) (Buswell et al., 2002; Carrió and Villaverde, 2003; Cheng et al., 1981). To derive active protein from these IBs, protein refolding is required and often provides low yields. Considerable amounts of work have looked at the effect of factors influencing refolding (Armstrong et al., 1999; De-Bernardez Clark et al., 1998; Dong et al., 2004; Katoh et al., 1999; Lee et al., 2002), but few have looked at the influence of upstream steps upon protein refolding (Batas et al., 1999; Buswell et al., 2002; Maachupalli-Reddy et al., 1997). Contaminating cell debris has been shown to increase the degree of proteolysis of IB (presumably through the presence of proteolytic enzymes) (Wong et al., 1996), affect the yield from refolding (Georgiou and Valax, 1999; Maachupalli-Reddy et al., 1997) and may present an extra burden on downstream processing steps (Thatcher et al., 1996).

To understand the nature and levels of contaminants present in purified IBs, it is critical to analyse the process steps used to purify them (Figure 4.1).

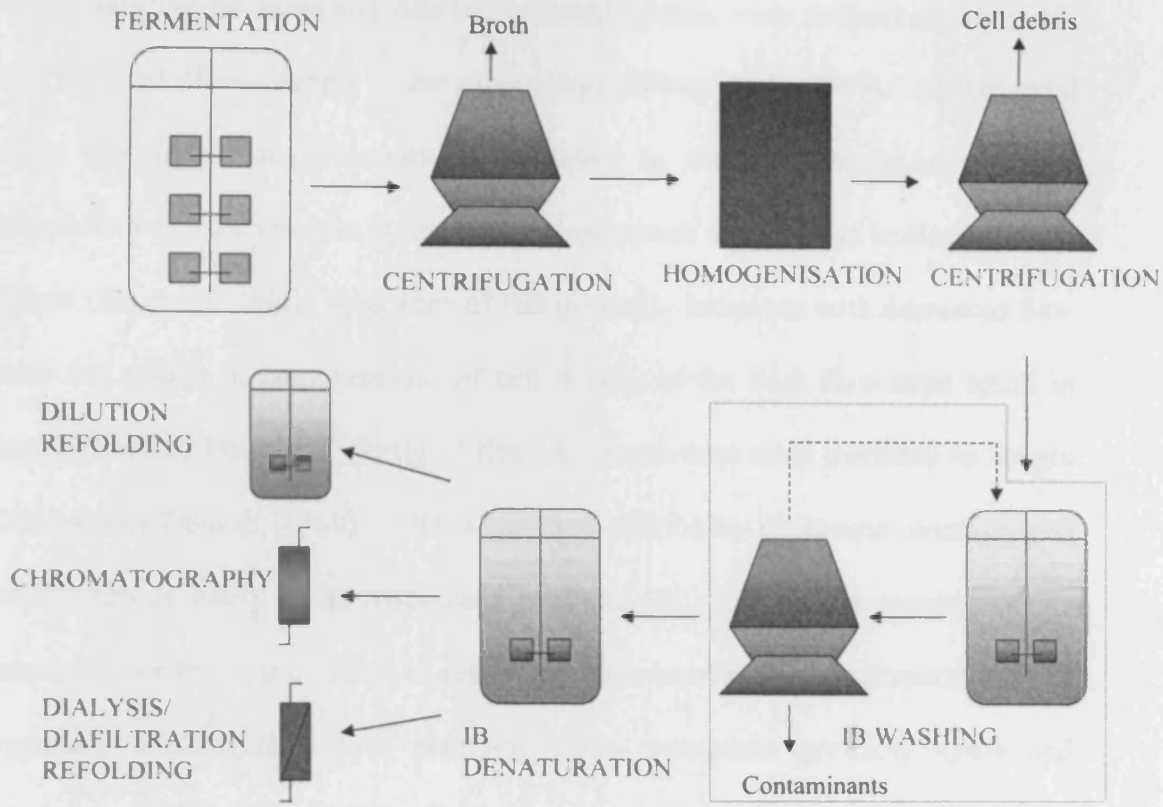


Figure 4.1: Inclusion body purification process. IB washing step is indicated in the boxed area. Several wash steps are usually used, each consisting of a re-suspension (wash) and a centrifugation step. Cells harvest, IB harvest and IB washing can alternatively be achieved by microfiltration (MF). Based upon (Titchener Hooker et al., 1991)

After fermentation, IBs are commonly released from the cells by homogenisation (Fischer et al., 1993; Titchener Hooker et al., 1991; Wong et al., 1996; Wong et al., 1997). IBs are more resistant to shear than other cellular structures (Thatcher et al., 1996). This was evidenced in studies where IB size did not appear to be affected by the number of homogenisation passes, whilst cell debris size was affected considerably (Wong et al., 1997). Therefore, given sufficient passes through a homogeniser a suspension containing a mixture of small cell debris particles and larger IB particles should be produced. Analysis of cell debris shows that it is heterogeneous in both size and density, whilst IBs are larger, more dense, and highly refractile, with greater homogeneity in size and optical properties (Jin et al., 1994). This difference in size and density can be exploited

to separate the IB from cell debris by centrifugation, with sedimentation of IB, and removal of cell debris in the supernatant (Wong et al., 1997). At industrial scale this separation is commonly achieved in disc-stack centrifuges where adjustment of flow rate can favour the sedimentation of inclusion bodies over cell debris (Jin et al., 1994). Recovery of IBs generally increases with decreased flow rate, but results in poor removal of cell debris, whilst high flow rates result in lower recovery but higher purity of IBs. A compromise must therefore be sought (Hoare and Dunnill, 1989). This separation will be by no means absolute, and cell debris is likely to be associated with the IB. Cell debris consists of the insoluble matter in the cells i.e. cell walls, membranes etc. IB preparations are typically contaminated with peptidoglycans, membrane proteins, lipids and nucleic acids (Thatcher, 1990). To remove these contaminants washing steps are usually used. These washing steps typically employ reagents, such as urea, at low concentrations to remove non-specifically bound proteins, and detergents to remove membrane and membrane proteins (Batas et al., 1999; Thatcher et al., 1996). Such steps can be followed with washes in buffer to remove the reagents. Laboratory scale purification procedures tend to employ one or more enzymatic steps, for example: DNAse to remove DNA and lysozyme to remove cell wall debris (Buswell et al., 2002).

This chapter observes how purification steps impact upon the relative levels of three major contaminants; protein, DNA and lipid, and how these in turn impact upon refolding yields. This contrasts with previous work, where single contaminants were refolded with lysozyme (Maachupalli-Reddy et al., 1997). Their study provided a guide to the effect of individual components on refolding,

but failed to deal with the effect of components in combination and with a real mixture of components, i.e. those present in a process stream. This chapter seeks to address this. Other studies have looked at the effect of washing stages but have failed to quantify the levels of contaminants removed (Batas et al., 1999).

This chapter concentrates on two key steps in IB processing: centrifugation and washing and observes their effect on the levels of IB contaminants and refolding yield, using trypsinogen in both pure and inclusion body forms. This chapter also details a series of studies to determine optimised refolding conditions for trypsinogen prior to studying the effect of contaminants. The reason for this was two fold: First to see if optimised conditions reduce the effect of contaminants, secondly to improve yields such that any perturbation in refolding efficiency can be more easily seen.

Trypsin is an industrially important enzyme. It is used in cell culture for the detachment of surface adherent cells, in vaccine manufacture, in the production of influenza virus and in the biopharmaceutical industry in the processing of hormones and various other proteins (Hohenblum et al., 2004).

Refolding was performed in batch mode since this allows a multitude of conditions to be tested and replicated over a short period of time. Batch refolding, however, represents a worst-case scenario, with poor mixing and rapid establishment of final protein concentration, resulting in lower yields. Optimal conditions found by batch methods may be further improved by use of fed-batch dilution methods as discussed in Chapter 2 and 3. Batch refolding hence provides an initial rapid screening method to identify optimal refolding conditions and conditions, which perturb refolding.

4.2 Materials and Methods:

4.2.1 Materials & Equipment

Bovine trypsinogen (1x crystallised) was purchased from Worthington Biochemical (Lakewood, NJ, USA). All other chemicals were purchased from Sigma Aldrich (Poole, Dorset, UK), with the exception of tryptone and yeast extract which were purchased from Oxoid Ltd (Basingstoke, Hampshire, UK) and kanimycin monosulfate, which was purchased from Fisher (Loughborough, Leicestershire, UK). ELTRPI was kindly supplied by Eli-Lilly (Indianapolis, IN, USA).

Centrifugation of samples (with the exception of the cell suspension) was achieved in an 4810 R centrifuge (Eppendorf, Hamburg, Germany). Cell suspension was centrifuged using a J2-M1 centrifuge (Beckman-Coulter, Fullerton, CA, USA) with JA10 rotor. Sonication was achieved using a MSE Soniprep 150 (23 KHz generator) (MSE, London, UK) equipped with a process timer. 96 well plates were assayed using the Saffire^2 plate reader (Tecan, Mannedorf, Switzerland) using Xfluor4saffireII software. Other spectrophotometric measurements were performed on a Genesys6 spectrophotometer (Thermo-Spectronic, Rochester, NY, USA).

4.2.2 Cell Growth

4.2.2.1 Growing ELTRP I on TB Media

ELTRP I from glycerol stock cells were grown on TB (terrific broth) media (12g/L tryptone, 24 g/L yeast extract, glycerol 0.4 % (v/v), 0.0170 M KH_2PO_4, 0.072 M K_2HPO_4) containing tetracycline (10 μg/mL) overnight at 32 °C. This stock was then used to inoculate a set shake flasks (5x 1 L) containing sterile TB

media with 10 μg/mL tetracycline. Cells were grown for 1h at 32 °C, and then at 37 °C until the end of exponential growth. Growth was monitored at hourly intervals. Confirmation of expression was achieved by SDS-PAGE. Homogenate, supernatant and pellet were prepared from a 100 mg/mL cell homogenate: Cells were resuspended at 100 mg/mL in homogenisation buffer and homogenised for 5 passes at 550 bar in the Lab 40 Homogeniser (APV Invensys, London, UK). Homogenate was spun down at 14000 rpm for 10 mins. Supernatant was removed for analysis. The pellet was resupended in the same volume of homogenisation buffer. 5-fold diluted samples of homogenate, supernatant and pellet were analysed by SDS PAGE.

4.2.2.2 Growing XL10 Cells on TB

XL10 cells (Kan+) were grown overnight on TB media (containing 50μg/mL kanimycin). The next day this stock was used as a 1/10 innoculum for 5 x 1L shake flasks (each final volume : 200 mL containing TB with 50 μg/mL kanimycin), this was grown for 7h at 37 °C. Growth was monitored by OD 600 nm at hourly intervals after inoculation from overnight stocks.

4.2.3 Assay Techniques

4.2.3.1 Assay of Trypsinogen Activity

Standards were prepared for each of the refolding conditions by diluting the appropriate denaturing solution without tryspinogen 10-fold in the appropriate buffer. Native trypsinogen was added to a final concentration of 0.1 mg/mL and then diluted to form a standard curve. The 0.1 mg/mL sample represents a 100 % yield, and enables an assessment of the amount of protein refolded.

100 µL of sample or standard was placed in the wells of a 96-well flat-bottomed plate and activated with 100 µL of 0.06 mg/mL enterokinase (in 20 mM $CaCl_2$, 40 mM Tris pH 8) by incubation for 2h at room temperature. 120 µL of activated solution was diluted with 80 µL of 1 mM N α-benzoyl-L-arginine 4-nitroanilide hydrochloride (L-BAPNA) (in 2.5% v/v dimethylformamide (DMF) in 20 mM $CaCl_2$, 40 mM Tris pH 8) and the change in absorbance at 405 nm was taken for 25 cycles at 20 second intervals (10 mins) and the rate of change in absorbance calculated. The amount of refolded trypsinogen was calculated from the standard curve.

4.2.3.2 SDS PAGE.

Gels were cast in the PROTEAN II system (Bio-Rad, Hercules, California, USA) composed of separating gel (12.6% acrylamide/bis-acrylamide (37:1), 0.375 M Tris, 1% SDS pH 8.8) and stacking gel (6% (w/v) acrylamide/bisacylamide (37:1), 0.125 M Tris, 1% (w/v) SDS, pH 6.8). Gels were set using 17.5 µL N,N,N',N'-Tetramethylethylenediamine (TEMED) and 100 µL 10 % (w/v) ammonium persulfate (APS) per 10 mL of gel. Samples were denatured and reduced by 1:1 dilution in 2x Laemmli sample buffer (4% (w/v) SDS, 20% glycerol (v/v), 10% (v/v) 2-mercaptoethanol, 0.004% (w/v) bromophenol blue and 0.125 M Tris HCl, pH 6.8) and were further denatured by boiling in water for 5 mins. Ladder (5 µL) (Precision plus protein standards, Biorad, Hercules, CA, USA)

Gel capturing was achieved on Gel Doc It imaging system (UVP, Upland, CA, USA) and analysed using Labworks software (UVP, Upland, CA, USA).

4.2.3.3 DNA Assay

Pre-sonicated samples (100 µL) were diluted 9-fold in homogenisation buffer (800 µl). Samples for standard curves were created as follows: Each diluted sonicate (450 µL) was further diluted 1:1 with 1mg/mL DNAse (450 µL) (in DNAse buffer (40 mM Tris, 6 mM $MgCl_2$, 2 mM $CaCl_2$) and left for 1 h to ensure DNA digestion. The DNAse was then inactivated by the addition of 100 mM EDTA (pH 8) (450 µL). 450 µl of this mixture was then used to dilute 50 µl of DNA stock (0.0725 mg/mL) (in TE buffer (10 mM Tris, 1 mM EDTA pH 7.4)) to produce the uppermost concentration of the standard curve. The remaining 900 µL of the mixture was diluted with TE buffer (100 µl)) to produce the diluent for the serial dilution.

Samples for measurement were generated as follows: each sonicate (450 µL)) was diluted 2-fold in inactivated DNAse (900 µL) (1 part 1mg/mL DNAseI (in DNAse buffer) 1 part 100 mM EDTA pH 8). This was further diluted by addition of TE buffer 150 µL).

Samples and standards were diluted 16.5 fold in TE buffer. 50 µL aliquots were assayed by addition of 50 µL of PicoGreen reagent (Invitrogen, Paisley, UK) (400x diluted in TE buffer) on a 96-well plate. Measurement of fluorescence intensity at 520 nm after excitation at 480 nm was used assess DNA levels. Amounts of DNA were calculated from the appropriate standard curve of that sample.

4.2.3.4 Lipid Assay

The lipid assay used was based upon a previous method (Izard and Limberger, 2003). A standard curve was created of glyceryl trioleate in 100 % chloroform. Aliquots of standards (100 µL) were placed into test tubes and the chloroform

removed by evaporation under a nitrogen stream, and replaced with 100 µL homogenisation buffer. 2 mL of 18 M sulphuric acid was added to 100 µL of each sample and each pre-prepared standard. These samples were then placed in a water bath at 100°C for 15 minutes. Samples were allowed to cool at room temperature for 5 minutes. To the cooled samples 5 mL of vanillin reagent was added (7.88 mM vanillin, in 68% (v/v) phosphoric acid). Samples were mixed and left at 37 °C for 15 minutes to allow the reaction to go to completion. Samples were assayed at 530 nm and concentrations calculated from the standard curve using a sample prepared with homogenisation buffer as a blank.

4.2.3.5 Protein Assay

Pre-sonicated samples were diluted 5-fold in homogenisation buffer. A standard curve was created using Bovine serum albumin (BSA) (in homogenisation buffer). BSA concentration was measured using an extinction coefficient of 0.667 mg^{-1} mL cm^{-1} (Peters, 1975).

Assay reagent was prepared using materials supplied in the BCA protein determination kit (Sigma Aldrich, Dorset, UK) by mixing 1 part 4% copper (II) sulfate pentahydrate with 50 parts of BCA reagent (bicinchoninic acid, sodium carbonate, sodium tartrate, and sodium bicarbonate in 0.1 N NaOH). Samples and standards were assayed in triplicate by dilution of 20 µL aliquots in assay reagent (200 µL). Samples were incubated at 37 °C for 30 minutes and then assayed at 562 nm.

4.2.4 Optimisation of Trypsinogen Refolding

4.2.4.1 Trypsinogen Denaturation

Trypsinogen was denatured at 1 mg/mL in 5.5 M Urea, 100 mM cysteine, 10 mM EDTA pH 9.5 (at 20 °C) for 2h at room temperature.

4.2.4.2 Effect of Various Refolding Cofactors

29 refold buffers were formulated as described in Table 4.1:

Cofactor	-	A	B	C	D	E	F	G
-	- (1)							
A	A (2)		AB (9)	AC (10)	AD (11)	AE (12)	AF (13)	AG (14)
B	B (3)			BC (15)	BD (16)	BE (17)	BF (18)	BG (19)
C	C (4)				CD (20)	CE (21)	CF (22)	CG (23)
D	D (5)					DE (24)	DF (25)	DG (26)
E	E (6)						EF (27)	EG (28)
F	F (7)							FG (29)
G	G (8)							

Table 4.1: Buffer preparation key for factorial experiments. 29 buffers were prepared. Each buffer contained 5mM Tris, 3 mM cysteine, 1 mM cystine pH 8.55. Buffer additives were as follows (-) No additives, (A) 0.25 M Sucrose, (B) 0.5 M Glucose, (C) 0.022 M Glycerol, (D) 0.023 M PEG 300, (E) PEG 3350 0.002 M, (F): 2-pentanol 0.011 M, (G) Cyclohexanol 0.09 M.

Trypsinogen was denatured as described in section 4.2.4.1 Refolding was achieved by 10-fold dilution of denatured protein in each of the buffers in Table 4.1 at 4 °C in duplicate, and left overnight (16h) at 4 °C to allow the refolding to go to completion. To calculate yields specific activity in each buffer was used to estimate refolding yield using the assay described earlier in section 4.2.3.1.

A factorial analysis of the data was undertaken. The experiment was organised as a series of 2^2 factorial studies. An estimate of individual effects was calculated in a similar fashion to normal factorial experiments as the difference between the mean yield when the factor was present (in this case the factor on its own and in combination with other factors) subtracted from the mean yield of when the factor was absent (other factors individually and the control). Interactions were calculated from a smaller pool of data: the control (no factors added), each factor individually and the combination of the two factors.

The significance of effects was estimated from pure error of replicates as follows. The sample variance of each combination was first calculated.

$$s_i^2 = \frac{\Sigma_{j=1}^{2}(Y_{ij} - \bar{Y}_i)^2}{n-1} \qquad (4.1)$$

where calculation is achieved from the sum of squares of differences between observed response (Y_{ij}) and the mean of this response, divided by the number observations (n) minus 1.

The average of the sample variances for each of the factor combinations used to calculate an effect, were then computed.

$$s^2 = \frac{\Sigma_{i=1}^{X} s_i^2}{X} \qquad (4.2)$$

Where X is the number of factor combinations used to measure the effect of a factor. S_{eff} is then calculated accordingly (4.3):

$$s_{eff} = \frac{s}{\sqrt{\frac{X}{4}r}} \qquad (4.3)$$

Where r is the number of replicates used. Confidence intervals are then calculated as described earlier (3.2.8) where $\nu = X$.

4.2.4.3 Factorial Experiment to Determine Importance of Various Factors Upon Refolding Yield

Five factors were tested for their effects upon refolding yield. Their levels are detailed in Table 4.2:

Factor	Upper	Lower
Buffer	5 mM Borate pH 8.85 (at 20 °C)	5 mM Tris pH 8.55 (at 20 °C)
Redox	14.1:1	4.44:1
Salt	264 NaCl/ 11KCl	0
Detergent	0.5% Triton X-100	0
Glucose	0.5M	0

Table 4.2: Levels for factorial experiment

16 buffers were prepared as detailed in Table 4.3:

Buffer	Species at 5mM	$CaCl_2$ (mM)	Cysteine (mM)	Cystine (mM)	Detergent (% v/v)	Salt (NaCl/KCl) (mM)	Glucose
1	Tris	50	0	2.5	0	0	0.5
2	Borate	50	0	2.5	0	0	0
3	Tris	50	1	3	0	0	0
4	Borate	50	1	3	0	0	0.5
5	Tris	50	0	2.5	0.5	0	0.5
6	Borate	50	0	2.5	0.5	0	0
7	Tris	50	1	3	0.5	0	0
8	Borate	50	1	3	0.5	0	0.5
9	Tris	50	0	2.5	0	264/11	0.5
10	Borate	50	0	2.5	0	264/11	0
11	Tris	50	1	3	0	264/11	0
12	Borate	50	1	3	0	264/11	0.5
13	Tris	50	0	2.5	0.5	264/11	0.5
14	Borate	50	0	2.5	0.5	264/11	0
15	Tris	50	1	3	0.5	264/11	0
16	Borate	50	1	3	0.5	264/11	0.5

Table 4.3: Buffer for 2^5 half factorial experiment

Refolding was achieved by 10-fold dilution of denatured protein (prepared as detailed in section 4.2.4.1) in each buffer (Table 4.3) at 4 °C in quadruplicate, and left overnight (16h) under the same conditions to allow the refolding to go to

completion. Samples and appropriate standards were prepared and assayed as described. Complete standard curves were made for each of the conditions. Assays were performed as described in section 4.2.3.1.

Effects were calculated as described previously. First the sample variance of each response was calculated as described in equation (4.1). These are then averaged to give an estimate of the variance for the complete set of data (s^2).

$$s^2 = \frac{\Sigma_{i=1}^{2k-1} s_i^2}{2^{k-1}} \qquad (4.4)$$

Where 2^{k-1}, gives the number of factor combinations in a half factorial experiment of k factors. S_{eff} is then calculated accordingly (4.5):

$$s_{eff} = \frac{s}{\sqrt{2^{k-1-2} r}} \qquad (4.5)$$

Where r is the number of replicates used. Confidence intervals are then calculated as described in section 3.2.8 where $\nu = 2^{k-1}$.

4.2.4.4 Effect of Salt Concentration

The effect of the NaCl/KCl ratio was tested. Trypsinogen was denatured as described in section 4.2.4.1. Refolding was achieved by 10-fold dilution of denatured protein in two buffers (20:1 ratio buffer) 5mM Tris, 0.5 M glucose, 50 mM $CaCl_2$, 2.5 mM cystine, 200 mM NaCl, 10 mM KCl ph 8.55 (at 20°C) or (24:1 buffer) 5mM Tris, 0.5 M glucose, 50 mM $CaCl_2$, 2.5 mM cystine, 200 mM NaCl, 8.3 mM KCl pH 8.55 (at 20°C) at 4 °C in triplicate, and left overnight (16h) under the same conditions to allow the refolding to go to completion. Samples and appropriate standards were prepared and assayed as described in section 4.2.3.1.

To test the effect of salt concentration 5 buffers were prepared as described in Table 4.4. Denatured protein was refolded by 10-fold dilution in each of the refold buffers at 4°C in triplicate, and left overnight (16h) under the same conditions to allow the refolding to go to completion. Samples and appropriate standards were prepared and assayed as described in section 4.2.3.1.

Buffer	Tris (mM)	$CaCl_2$ (mM)	Cystine (mM)	Glucose (M)	NaCl (mM) (final)	KCl (mM) (final)	PH (at 20 °C)
1	5	50	2.5	0.5	0	0	8.55
2	5	50	2.5	0.5	236 (212)	11.8 (10.6)	8.55
3	5	50	2.5	0.5	500 (450)	25 (22.5)	8.55
4	5	50	2.5	0.5	750 (675)	37.5 (33.75)	8.55
5	5	50	2.5	0.5	1500 (1350)	75 (67.5)	8.55

Table 4.4: Buffers used when measuring the effects of salts on refolding yields.

4.2.5 Effect of Processing Steps Upon IB Purity and Refolding Yields

4.2.5.1 Modelling Disc Stack Centrifuge Efficiency in a Lab Scale Centrifuge

Assessment of the effect of centrifugation efficiency in an industrial disc stack centrifuge (CSA 8-06-476, Westafalia, Oelde, Germany) was modelled using an

ultrascale-down model mimic based upon a laboratory centrifugation. This model used the equations described previously (Boychyn et al., 2004).

A rotor speed of 6000 rpm in the lab centrifuge was used throughout. This was used to produce spin times equivalent to the following disc stack centrifuge flow rates 210, 479, and 980 Lh^{-1} (required spin times of 9, 4 and 2 mins respectively at 6000 rpm).

4.2.5.2 Determining the Effect of Centrifugation Efficiency on Refolding and Contaminant Amounts

ELTRPI and XL10 cells were grown up as described in section 4.2.2. Cells were spun down for 30 mins at 10000 rpm. Cells were resuspended to a final concentration of 100 mg/mL in homogenisation buffer. Cells were then subject to 5 passes at 550 bar in a Lab 40 homogeniser. 2 mL aliquots each of XL10, induced ELTRP homogenate and uninduced ELTRP I sonicate were centrifuged in the lab scale centrifuge for 10 mins at 14000 rpm. Homogenate, resuspended pellet and supernatant for each of the suspensions were run on SDS-PAGE. The percentage trypsinogen in the induced ELTRP pellet was approximated by densitometry. The protein concentrations of these pellets were measured by BCA assay (section 4.2.3.5). Densitometry and protein concentration of ELTRPI pellet were used to calculate the final volume of resuspended pellet that gave approximately 1mg/mL trypsinogen. This final volume was used for all ELTRPI resuspensions/solublised under all conditions. This was used to calculate the final resuspended volume of XL10 pellet such that it contained a comparable amount of non-trypsinogen protein as ELTRPI (calculated as the total amount of protein in the ELTRPI pellet –1 mg/mL). This final volume was used for all XL10 resuspensions/solubilisations under all conditions.

2 mL aliquots of XL10 and ELTRPI homogenate were spun down for 2, 4 and 9 minutes at 6000rpm to mimic flow rates of 210, 479 and 980 Lh^{-1} (Q/Σ= 23.08, 52.54, and 107.37 $x10^{-9}$ m/s) respectively, and at 14000 rpm for 10 minutes to allow complete settling of cell debris and inclusion bodies (equivalent to a flow rate of zero). Supernatant was removed and OD taken at 600 nm and 420nm, to establish the OD_{600nm}/OD_{420nm} ratio, which provides an approximation of relative levels of IB to cell debris. The assumption behind this ratio is that IB particles scatter light more efficiently at 600 nm compared to the degree of scatter at 420 nm than do cell debris particles (Jin et al., 1994). OD was corrected against the fully spun down suspension (Flow rate 0), hence fully spun down suspension had a ratio of 0. Pellets were stored at 4 °C prior to use.

Pellets generated under each of the centrifugation conditions were resuspended in homogenisation buffer, by sonication (3x 10s at 10 μ amplitude) to the final volume as calculated earlier. DNA, lipid and protein concentrations were measured as described in sections 4.2.3.3, 4.2.3.4 and 4.2.3.5 respectively. Assessment of the purity of centrifuged pellets was by SDS-PAGE.

ELTRP I pellets from each of the centrifugation conditions were resuspended and solubilised in solublisation buffer at a final concentration of 5.5M urea, 100 mM cysteine, 10 mM EDTA pH 9.5. XL10 pellets produced under each of the centrifugation conditions were resuspended and solubilised in solublisation buffer at a final concentration of 5.5 M urea, 100 mM cysteine, 10 mM EDTA pH 9.5 and separately in denatured trypsinogen to a final concentration of 1mg/mL

trypsinogen in 5.5 M urea, 100 mM cysteine, 10 mM EDTA pH 9.5. The former was used for standard curves and the later for refolding. Sonication (3 x 10 s at 10μ amplitude), was used to resuspend pellets to aid solublisation. Solubilised samples were incubated at ambient temperature for 2h.

Solubilised samples were refolded by 10-fold dilution in triplicate in two buffers: Buffer (1): 5 mM Tris, 50 mM $CaCl_2$, 3 mM cysteine, 1 mM cystine pH 8.55 (at 20°C) and Buffer (2) 5 mM Tris, 50 mM $CaCl_2$, 0.5 M NaCl, 0.5 M Glucose, 2.5 mM cystine pH 8.55 (at 20°C) overnight at 4 °C (16h) to allow the reaction to go to completion. 10-fold diluted XL10 samples were used to form standard curves for each of the centrifuge conditions, refolding yields for ELTRPI refolds and XL10 pellets containing tryspinogen were assessed from these standard curves.

4.2.5.3 Determining the Effect of Washing Steps on Levels of Contaminants and Refolding Yield

XL10 and ELTRP I cells were grown and homogenised as described in 4.2.5.2. Levels of trypsinogen in the pellet were assessed as described in the previous section. Final resuspension volume for each cell type was calculated as described as in the previous section. These final volumes were used for washes, resuspension and solubilisation.

2mL aliquots of homogenised cells were centrifuged at 14000 rpm for 10 minutes and supernatant removed. These pellets were used to observe the effect of washing steps in a factorial study.

This study was organised in a factorial fashion using three factors: Triton, Urea and sonication for ELTRP I homogenate and just urea and Triton for XL10 homogenate. Sonication was used as an additional step to see if maximising surface area for washing could improve washing efficiency and improve refolding

yields. This step would not necessarily be applicable to industry, but may reveal the importance of particle size to washing efficiency.

A total of four wash stages were performed for each run, details of which are described below in Table 4.5 and 4.6. A wash involved resuspending the pellet in wash buffer with or without the application of sonication (3x 10s at 10μ amplitude) and then centrifuging the suspension (10 mins at 14000 rpm) to produce the washed pellet. The supernatant was then removed and the next wash started. All steps were performed on ice, and all samples were stored on ice throughout

ELTRP I

Run	Wash1	Sonication	Wash2	Sonication	Wash 3	Sonication	Wash
1	Triton	Yes	Urea	Yes	Homog	Yes	Homog
2	Triton	No	Urea	No	Homog	No	Homog
3	Triton	Yes	Homog	Yes	Homog	Yes	Homog
4	Triton	No	Homog	No	Homog	No	Homog
5	Homog	Yes	Urea	Yes	Homog	Yes	Homog
6	Homog	No	Urea	No	Homog	No	Homog
7	Homog	Yes	Homog	Yes	Homog	Yes	Homog
8	Homog	No	Homog	No	Homog	No	Homog
9	None	-	-	-	-	-	-

Table 4.5: Wash steps performed on ELTRP I cells. Where: Urea: 1M Urea, 50 mM Tris, 50 mM NaCl pH 8.8. Triton: 0.5% v/v Triton X-100, 50 mM Tris, 50 mM NaCl pH 8.8. Homog: 50 mM Tris, 50 mM NaCl pH 8.8.

XL10 cells

Run	Wash1	Sonication	Wash2	Sonication	Wash 3	Sonication	Wash
1	Triton	No	Urea	No	Homog	No	Homog
2	Homog	No	Urea	No	Homog	No	Homog
3	Triton	No	Homog	No	Homog	No	Homog
4	Homog	No	Homog	No	Homog	No	Homog

Table 4.6: Wash steps performed on XL10 cells.

Pellets for analysis of protein, DNA and lipid concentration were resuspended in homogenisation buffer, by sonication (3x 10s at 10μ amplitude) to the final volume calculated earlier. Protein, DNA and lipid concentrations were measured as described in 4.2.2.5, 4.2.2.3 and 4.2.2.4 respectively. An assessment of the purity of centrifuged pellets was by SDS-PAGE.

ELTRP I pellets were resuspended and solubilised in solublisation buffer at a final concentration of 5.5 M urea, 100 mM cysteine, 10 mM EDTA pH 9.5. XL10 pellets were resuspended and solubilised in solublisation buffer at a final concentration of 5.5M urea, 100 mM cysteine, 10 mM EDTA pH 9.5 and separately in trypsinogen final concentration 1mg/mL in 5.5 M urea, 100 m cysteine, 10 mM cysteine pH 9.5, the former being used for standard curves and the later for refolding. Sonication (3 x 10 s at 10 μ amplitude), was used to resuspend pellets to aid solublisation. Solubilised samples were incubated at ambient temperature for 2h.

Solubilised samples were refolded by 10-fold dilution in triplicate in two buffers: (buffer 1) 5 mM Tris, 50 mM $CaCl_2$, 3 mM cysteine, 1 mM cystine pH 9 (at 20°C) and 5 mM Tris, 50 mM $CaCl_2$, (buffer 2) 0.5 M NaCl, 0.5 M Glucose, 2.5 mM cystine pH 8.55 (at 20°C) overnight at 4 °C (16h) to allow the reaction to go to completion. 10-fold diluted XL10 were used to form standard curves for each of the conditions, refolding yields for ELTRPI refolds and XL10 pellets containing tryspinogen were assessed from these standard curves.

The significance of each effect was adjudged from pure error of assay replicates. Effects measured as being greater than this at a 95% confidence interval were assumed significant.

4.3 Results and Discussion

4.3.1 Optimisation of Trypinogen Refolding

Initial tests (results not shown) revealed that the conditions used in a previous study on trypsinogen refolding (Buswell et al., 2002), could not be improved by using high concentrations of oxidising agents in the refolding buffer, DTT as a reducing agent or high concentrations of urea. Results did reveal that best yields were achieved with buffer at 4 °C. The following tests were used to improve refolding yields further.

4.3.1.1 Effect of Various Refolding Cofactors

Several factors appear to improve trypsinogen refolding (Hibbert, 2004), but to date studies have only approximated effects by increases or decreases in activity.

A factorial analysis of the results from this experiment, produced the effects plot (Figure 4.2). It is clear from this effects plot that the only factor to have a

significant positive effect overall is glucose and for this reason the optimisation of tryspinogen refolding used glucose as the sole factor for further tests.

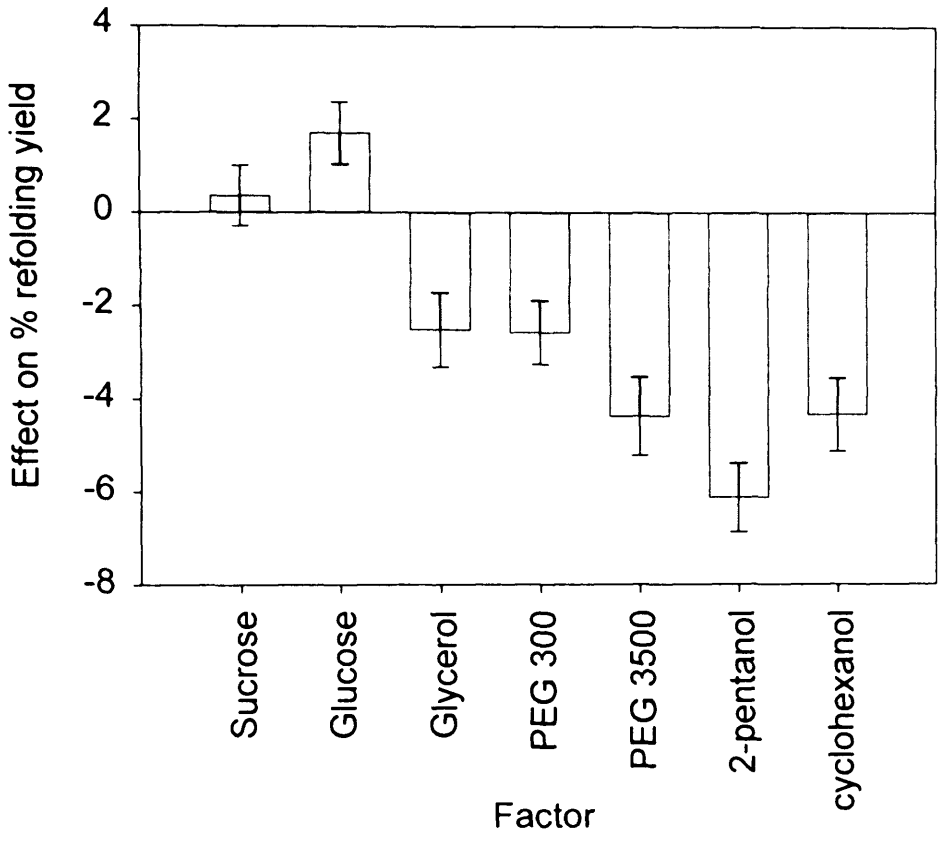


Figure 4.2: Effects plot for effect of cofactors. Plot indicating effect of single factors. Two factor interactions are not indicated since they did not reveal any positive interactions. Error bars indicate 95% confidence interval, bars larger than this interval are deemed significant.

4.3.1.2 Factorial Experiment to Determine Importance of Various Factors Upon Refolding Yield

It has been suggested (Willis et al., 2005) that addition of salts and detergents may improve refolding yields of serine proteases. A half-factorial experiment was conducted to understand the importance of various factors upon refolding yield. The levels of factors used in the following study are indicated in Table 4.2. A factorial analysis was conducted on the data to give the effects plot shown in Figure 4.3.

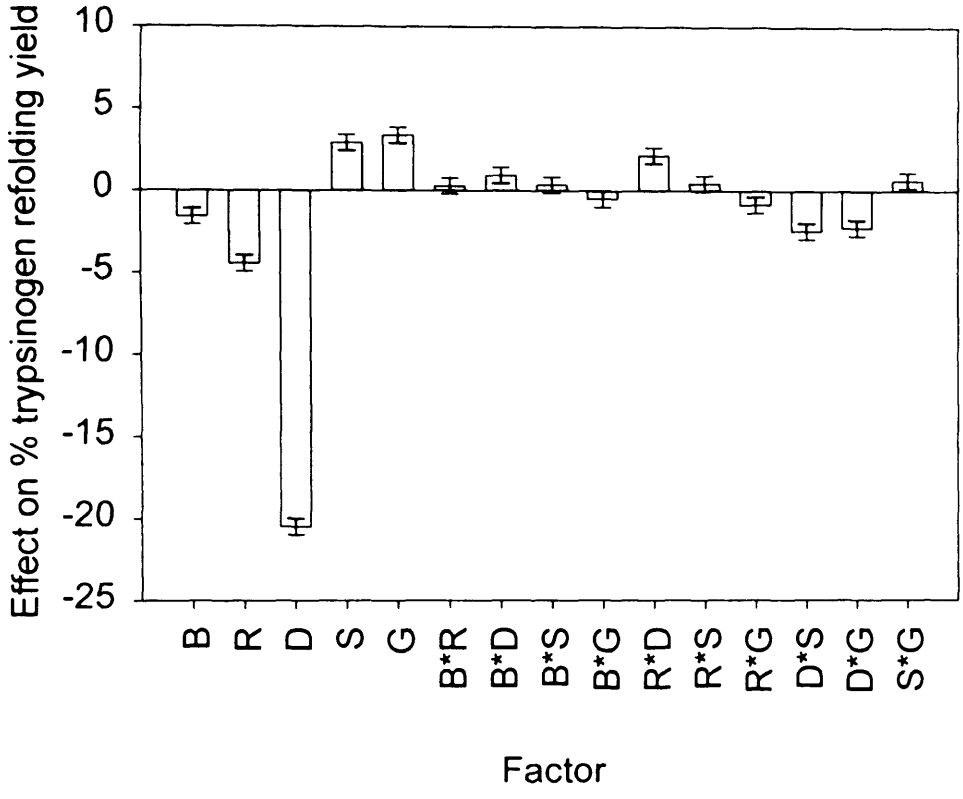


Figure 4.3: Effect of factors upon refolding yield. Where B= buffer, R= Redox ratio, D= Detergent, S= Salt, G= Glucose. Error bars indicate 95% confidence interval, effects larger than the error bars are deemed significant.

It is evident from this analysis that each of the factors tested appear to be significant as a single factor. In addition it appears there were significant interactions between detergent and each of the other factors tested (B*D, R*D, D*S, D*G), and a significant interaction between redox ratio and glucose.

The most significant effect appears to be that of detergent. It is thought that detergent may prevent aggregation through its chaotropic quality, whereby prevention of aggregation may permit higher refolding yields (De-Bernardez Clark, 1998). However it appears at least for trypsinogen, that it reduces yields considerably. This large negative effect results in the interactions observed

between this and other factors (Figure 4.3). The effect of each factor appears to be decreased in the presence of detergent, and is so strong that other factors are unable to compensate.

Glucose has a positive effect and also the addition of salts also appears to improve yields. Salts stabilise the native state and may improve the rate of renaturation and refolding yields. The effect of such salts needs further analysis as presented in the next section.

It appears that using a slightly lower redox ratio buffer improves yields, perhaps by virtue of providing a more balanced ratio of reduced to oxidised species, which may allow more efficient disulfide shuffling.

The choice of buffer had a significant effect on refolding yields. Buffers usually operate to approximately 1 pH unit, above and 1 unit below their pKa. Borate has a pKa of 9.23 (range 8.23-10.23) whilst Tris has a pKa of 8.06 (range 7-9) (Beynon and Easterby, 1996). It was therefore envisaged that use of borate buffer, which has a higher pH range than Tris would prove more suitable for buffering at pH 9, and therefore generate better yields. However, it appears that Tris is a better buffer additive since higher yields are achieved. It is unclear why this is the case though the effect is the smallest of all the single factor effects examined.

4.3.1.3 Effect of Salt Concentration

Further investigation of the effect of salt concentration was required to understand whether increases in salt concentration improve refolding. Comparison of ratios of NaCl: KCl, showed very little difference between yields. A 24:1 ratio gave 32.58% (± 0.45%), whilst a 20:1 ratio 32.42 % (± 2.67%). It is clear from Figure

4.4, that the optimum final salt concentrations are 0.45 M NaCl/ 22.5 mM KCl (0.5M NaCl/ 0.25 mM KCl in buffer).

It can be concluded from these studies that the optimised buffer should contain Tris, have a low redox ratio, not contain detergent, contain added salts and also glucose.

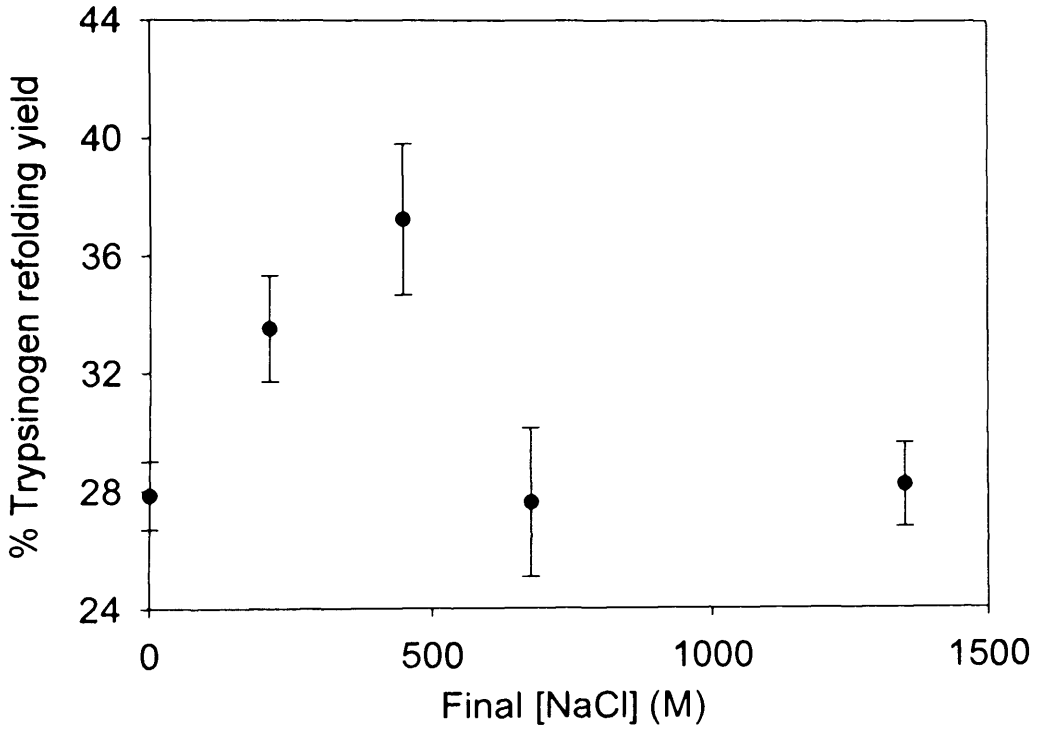


Figure 4.4: Effect of salt concentration on refolding yield. Where KCl concentration was 1/20th of NaCl. Standard deviation from triplicates.

4.3.2 Effect of Purification Steps on IB Purity and Refolding Yields

4.3.2.1 Determining the Effect of Centrifugation Efficiency on Refolding and Contaminant Amounts.

As discussed in the introduction to this section, centrifugation is usually used to separate IB from cell debris. A scale-down method mimicking different centrifugal flow rates was used to observe the effect of flow rates upon IB purity and refolding yields.

It can be seen from Figure 4.5 that the turbimetric ratio of the supernatant increases with an increase in Q/Σ. This increase in turbimetric ratio may indicate a greater presence of denser IB particles in the supernatant, at the higher Q/Σ. This means as flow rate increases the amount of IB retained in the solids underflow decreases. This result is to be expected. The higher the Q/Σ, the less the residence time available for separation and sedimentation. The time available may be sufficient for large dense particles to be separated, but may be insufficient for smaller less dense particles. As the Q/Σ increases so the residence time available falls, such that eventually all particles remain suspended and are unable to settle out.

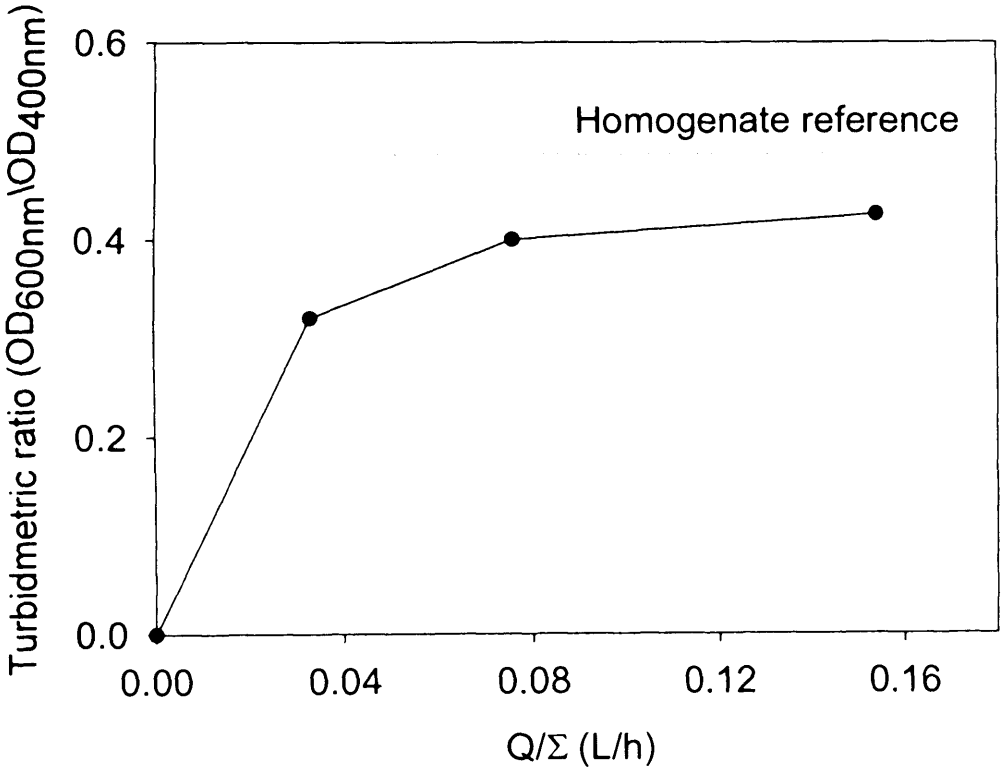
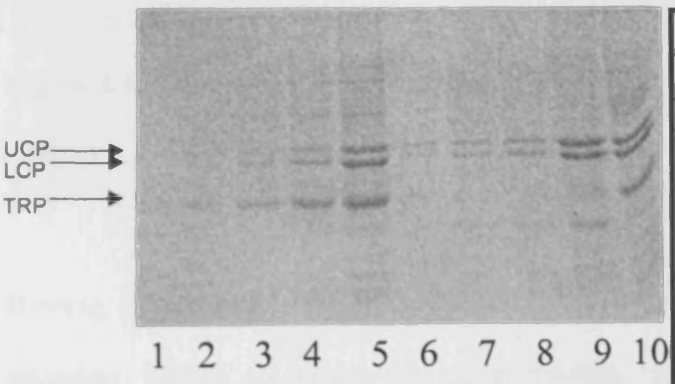


Figure 4.5: Effect of mimicked centrifuge flow rate on turbimetric ratio of ELTRPI supernatant.

As expected increases in Q/Σ reduces the total amount of inclusion body sedimented. It was unclear, from this initial analysis, how much of the recovered solids was trypsinogen, and how flow rate affected this. SDS-PAGE analysis (Figure 4.6), alone was not very informative, but it appears that the composition of IB, as approximated by densitometry (Figure 4.7), increases in trypsinogen purity as Q/Σ rises, but this is compromised by considerable losses in protein. Two major contaminant bands are evident and may correspond to 37 kDa and 40 kDa *E.coli* outer membrane proteins observed in previous studies (Jin et al., 1994). The presence of these bands in both IB expressing ELTRPI and non-inclusion body expressing XL10 strains strengthen this case, though the bands appear to be running at slightly higher apparent molecular weights.



Well	ID
1	Trypsinogen 0.1 mg/mL
2	ELTRP Q/Σ= 107.37 $1x10^{-9}$ m/s
3	ELTRP Q/Σ= 52.54 $1x10^{-9}$ m/s
4	ELTRP Q/Σ= 23.08 $1x10^{-9}$ m/s
5	ELTRP Q/Σ=0 m/s
6	XL10 Q/Σ= 107.37 $1x10^{-9}$ m/s
7	XL10 Q/Σ= 52.54 $1x10^{-9}$ m/s
8	XL10 Q/Σ= 23.08 $1x10^{-9}$ m/s
9	XL10 Q/Σ= 0 m/s
10	ELTRP Q/Σ= 0 m/s

Figure 4.6: SDS PAGE of centrifuged samples at various Q/Σ. Upper contaminating protein (UCP), lower contaminating protein (LCP), trypsinogen (TRP).

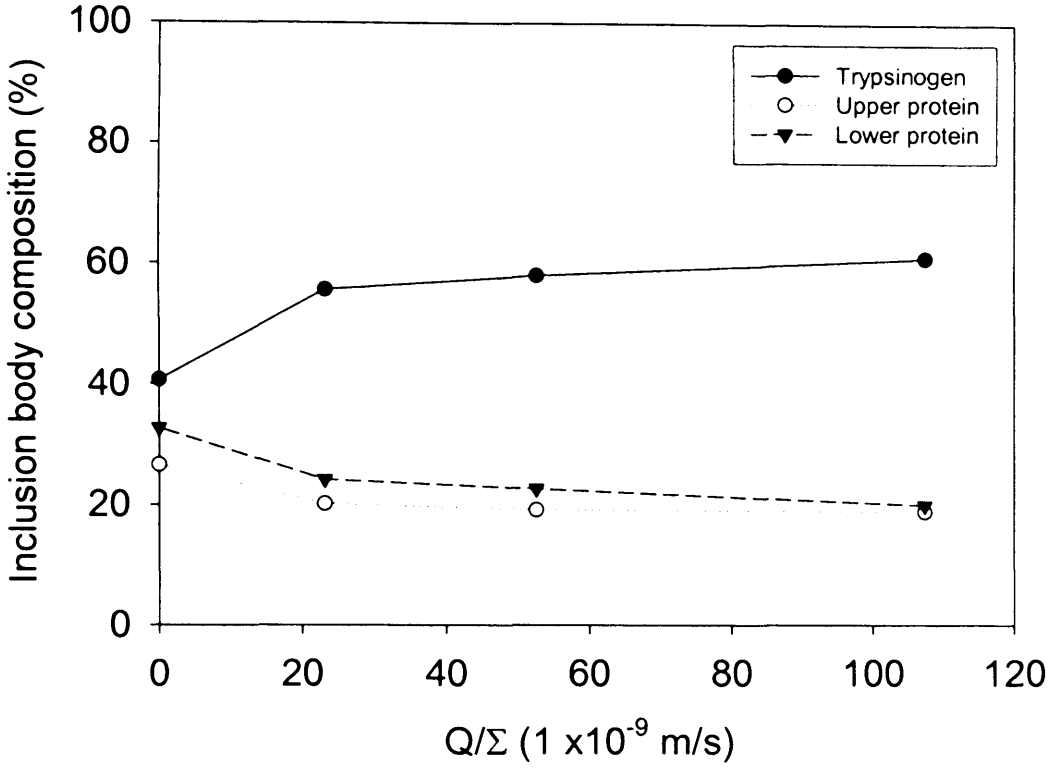


Figure 4.7: Effect of Q/Σ on composition of the three most abundant proteins in the solids fraction.

Having established the effect of centrifugal operation on protein composition attention turned to measuring total protein concentration, and to levels of contaminating DNA and lipid. From Figure 4.8 it appears as expected that levels of protein, DNA and lipid in the solids fraction decreased considerably with increasing Q/Σ, with removal of lipids, being more effective than for protein and DNA. Clearance of protein, lipid and DNA appeared to be slightly more efficient in the XL10 cells, which do not contain inclusion bodies, and may be because certain components adhere to IB more efficiently than they do to the cell debris. It is unclear why lipids are removed more efficiently than other components.

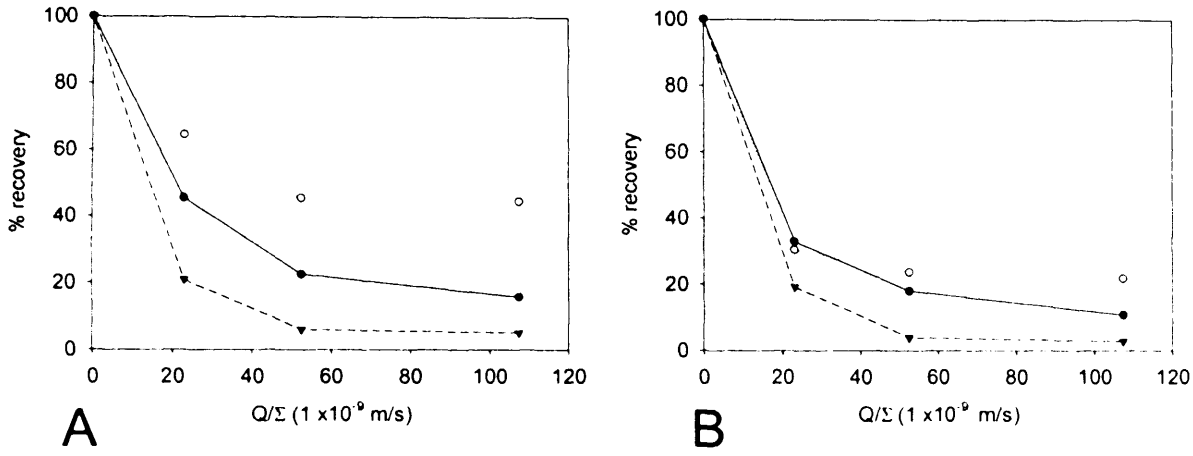


Figure 4.8: Effect of Q/Σ on levels of Protein (•), DNA (○) and Lipid (▼) in the solid phase. (A) ELTRPI cells. References levels for full sedimentation of solid phase: protein = 3429 μg/mL, lipid recovery = 829 μg/mL, DNA = 14 μg/mL. (B) XL10 cells . References levels for full sedimentation of solid phase: protein = 2197 μg/mL, lipid = 794 μg/mL, DNA = 10 μg/mL.

To establish the effect of contaminant removal achieved by centrifugation upon protein refolding solids fractions were solubilised to the same final volume. The protein concentration of the solubilised pellets will decrease with an increase in centrifuge flow rate, as protein is lost to supernatant. Refolding yields (in Figure 4.9) are hence based upon on the total amount of protein refolded.

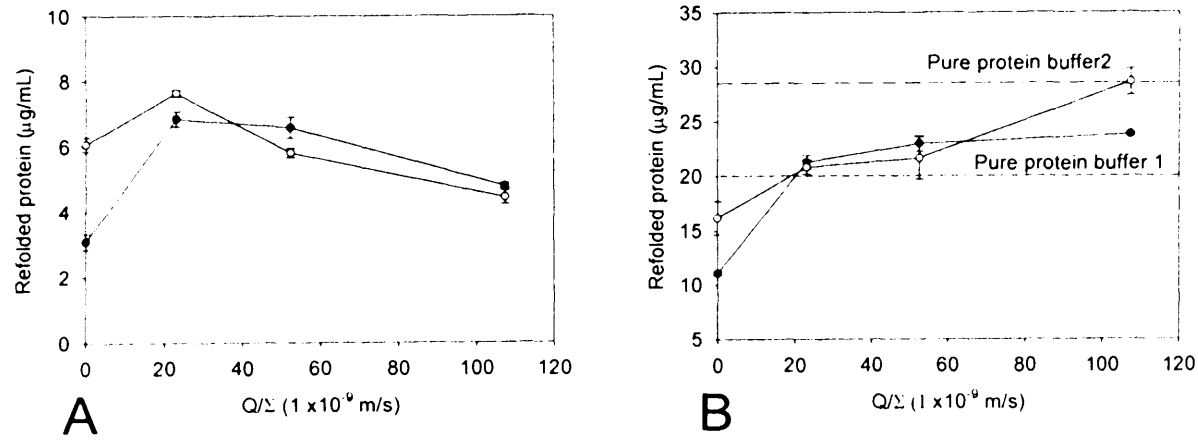


Figure 4.9: Effect of centrifugation flow rate on refolding yield (• buffer 1, ○ buffer 2) from (A) ELTRP (B).XL10 pellets solubilised with pure trypsinogen (final concentration 1 mg/mL). Trypsinogen yield in the absence of contaminants in Figure 4.9B for each buffer.

Refolding studies were performed on collected solid fractions at each of the centrifugation conditions using two buffers. Buffer 1 was that based upon previous studies (Buswell et al., 2002) and buffer 2 was that optimised in studies earlier in this chapter containing glucose, salt, and an optimised redox ratio.

Refolding from inclusion bodies results in lower yields than when refolding from pure protein in the presence of contaminants. In the case of ELTRPI cells the greatest amount of refolded protein was achieved by processing material separated at $Q/\Sigma = 23.08 \times 10^{-9}$ m/s. This Q/Σ appeared to give a considerably less contaminated inclusion body than the base case (maximal time for sedimentation), with slightly less purity in terms of protein than when refolding from solids generated at higher Q/Σ. The greater amount of IB recovered when compared to faster Q/Σ may have provided the comparatively higher yields. In addition the material separated at $Q/\Sigma = 23.08 \times 10^{-9}$ m/s appears to contain proportionately more lipid per protein and less DNA/ protein (Figure 4.10) than at higher flow rates. Lipids have been shown to increase yields of refolding proteins (Maachupalli-Reddy et al., 1997), and this may be the case here. Optimised buffer 2 appears to be less affected by the levels of contaminants present in the solids processesed at the base case. It appears the components within the optimised buffer 2 may mitigate the negative effects of contaminants allowing greater refolding yields.

Studies with XL10 provided an idea of the effect of the cell debris contaminants upon refolding of protein on pure tryspinogen. It appears that the level of contaminants present in a well-spun sample (base-case) suppresses refolding. It is likely that this is due to the high level of contaminating protein that may be co-

aggregating with the refolding protein preventing refolding. At $Q/\Sigma = 23.08 \times 10^{-9}$ m/s and above sufficient amounts of contaminants have been cleared to allow refolding yields, comparable with those where no contaminant is present. In the case of buffer 1 yields in the presence of contaminants ($Q/\Sigma = 23.08 - 107.37 \times 10^{-9}$ m/s) appear to be slightly higher than in their absence. Under these situations the levels of protein and DNA, shown previously to affect refolding yields (Maachupalli-Reddy et al., 1997) may be sufficiently low not to affect refolding yields, whilst contaminants such as lipids may still be at a sufficient concentration to improve refolding yield. This effect was absent for buffer 2 where maximal clearance of contaminants was essential to gain maximal yields.

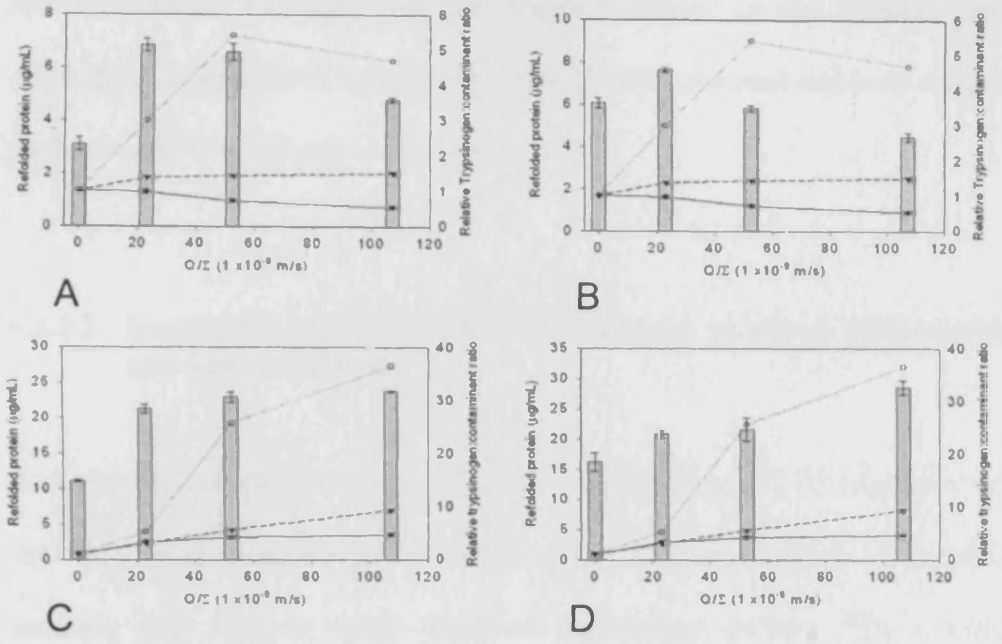


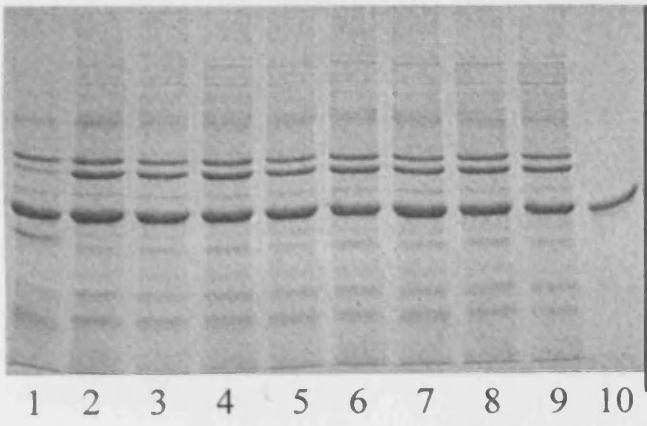
Figure 4.10: Effect of Q/Σ on approximate ratios of trypsinogen to contaminant levels and on the resultant and refolding yield. Relative trypsinogen: contaminants ratios are compared to the fully spun down pellet. Bars show refolding yield, ○= relative trypsinogen to lipid ratio,●= relative trypsinogen to DNA ratio, ▼=relative tryspinogen to contaminating protein.A) Refolding from ELTRP1 pellets using refold buffer 1, (B) Refolding from ELTRP1 pellets using refold buffer 2, (C) Refolding from pure tryspinogen in the presence of XL10 cell debris using refold buffer 1, (D) Refolding from pure tryspinogen in the presence of XL10 cell debris using refold buffer 2.

A critical observation from this study is that the ratio of contaminants to refolding protein is critical to the refolding yield achieved. It appears that refolding yield increases with decreasing contaminant. A critical level of contaminant is required before yields are affected negatively. Clearance of such an amount is required to maximise eventual refolding yields. It is also evident from figure 4.10 A-D, that levels of contaminating protein are particularly important to final refolding yields with levels of lipids having possible positive effects upon refolding. Clearly centrifugation has some effect in reducing the levels of contaminant associated with IB, but significant removal of contaminants is always at the expense of

losses in target material, and this appears critical to the eventual amounts refolded. A compromise exists where enough IB is retained and is of a significant purity to result in a good overall yield.

4.3.2.2 Determining the Effect of Washing Steps on Levels of Contaminants and Refolding Yield

Analysis of solid-phase fractions by SDS PAGE (Figure 4.11) appears to suggest that only sonication has a clear effect on tryspinogen purity. It is unclear if washing with urea or triton improves trypsinogen purity. The combination producing the cleanest IB appeared to be that produced by combining sonication, urea and Triton X washing (condition 1).



Lane	ID	% Tryp all	% Tryp (top3)
1	Condition 1	46.4	73.0
2	Condition 2	37.7	52.0
3	Condition 3	41.3	55.4
4	Condition 4	35.8	49.1
5	Condition 5	40.1	57.0
6	Condition 6	36.3	49.3
7	Condition 7	40.9	58.0
8	Condition 8	36.6	49.0
9	Unwashed	37.9	52.3
10	0.1 mg/mL trypsinogen	-	-

Figure 4.11: SDS-PAGE of IB washes. Where % Tryp all = calculated percentage trypsinogen of all the bands. % tryp (top 3)= calculated percentage of three most abundant bands.

Some loss of protein occurs in the washing stages especially for those employing sonication. The effect of different factors employed during washing upon total pellet protein concentration are revealed in the trends indicated in Figure 4.12. It

is clear for ELTRP I samples that sonication leads to the most efficient removal of protein (this effect is also apparent from the SDS-PAGE (Figure 4.11)). Urea appears also to have a major effect, which is to be expected as urea will solubilise some proteins removing them from the pellet and thereby reducing the protein concentration.

It appears that protein removal is more efficient from non-inclusion body containing cells (less insoluble matter to associate with). Both urea and Triton X have significant effects. Triton may remove protein associated with lipids such as membrane proteins. An interaction is present as results show the combined effect of Triton and urea to be less than the sum of their individual effects.

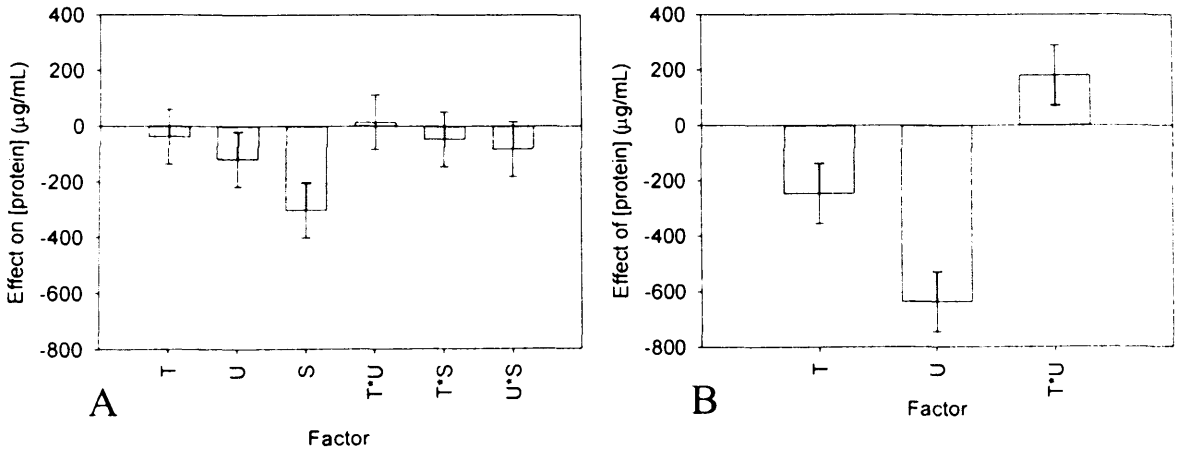


Figure 4.12: Effect of factors on protein concentration (A) ELTRPI pellets (B) XL10 pellets. Error bars are 95% confidence intervals. Bars larger that the error bars are deemed significant. Where T= effect of Triton X , U= effect of urea, S= effect of sonication, T*U= Triton/urea interaction, T*S= Triton/ sonication interaction, U*S= urea/sonication interaction.

The effects of factors upon lipid concentrations is shown in Figure 4.13. As may be expected Triton-X has a significant effect upon lipid levels, with similar effects for both ELTRPI and XL10 cells. Detergents are combined in washing steps to break up lipids, remove cell membranes and associated protein. Again sonication has a significant effect upon contaminant levels, presumably by generating a

greater surface area for washing whilst reducing cell debris size to facilitate removal. Urea only appeared to have a significant effect on XL10 lipid levels, where sonication was not used.

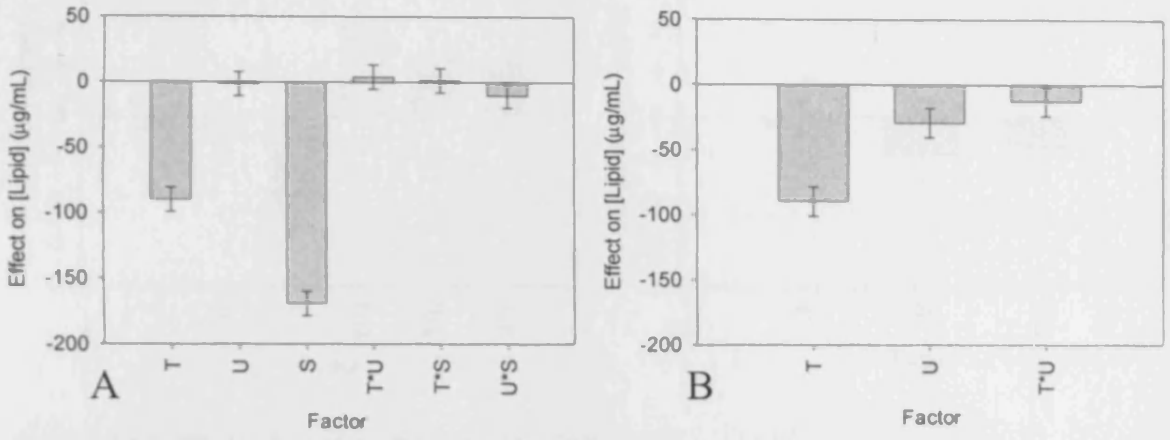


Figure 4.13: Effect of washing factors upon lipid levels: (A) ELTRPI (B) XL10. Error bars are 95% confidence intervals. Bars larger that the error bars are deemed significant.

As observed in the previous studies (Valax and Georgiou, 1993) DNA is a minor contaminant forming a very small percentage of the solid fraction. Figure 4.14 shows the effect of washing factors upon DNA levels within the pellet. It is clear for the ELTRPI pellet that sonication is most effective at removing DNA.

For XL10 cells in the absence of sonication urea has only a minor effect. Interactions are present between sonication and the other factors. Sonication appears to remove most if not all of the DNA (it was undetectable by assay), such that addition of other factors did not remove any more DNA. The positive interaction results as the effect of the sonication and other factors together having a lesser effect than the single factors in isolation. The interaction observed between Triton and urea is quite large, this appears to have come about as the effect of each factor is reduced in the others presence. When both are present DNA levels appear to be higher, but it is unclear why this is the case, and may be an artefact. Sonication has a considerable effect upon removal of each of the contaminants, with urea and Triton X impacting upon protein and lipid levels respectively.

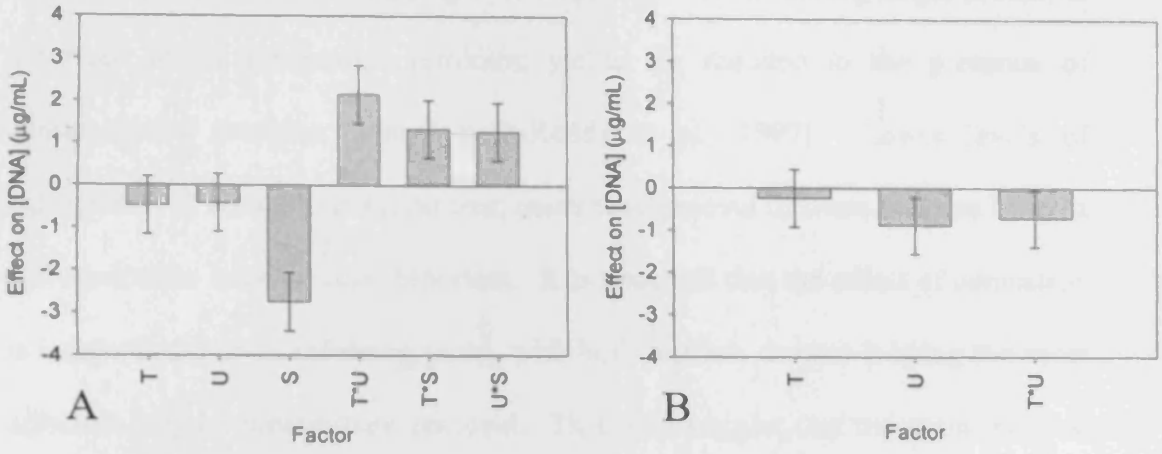


Figure 4.14: Effect of washing factors upon DNA levels.

Having determined the individual and combined effect of factors on contaminants the next stage of the research sought to establish the effect of washing steps upon the resulting protein refolding yield.

From Figure 4.15 (A-D) it appears that the most significant effect upon protein refolding yields from IB is Triton X. Triton X primarily appeared to effect the levels of lipids, affording efficient removal. This result suggests that removal of lipid might be critical to high refolding yields, but this is contrary to earlier results. The location of the Triton X step as the first stage of the experimental process may also be critical, since this stage has the greatest opportunity to remove material, it may remove proteolytic enzymes which are likely to affect refolding yields (Wong et al., 1997). Urea also appears to have an effect in the case of refolding using Buffer 1. As stated in the introduction the purpose of urea addition is to remove loosely associated protein from the IB. Removal of these proteinacious contaminants appears to have improved yields. The reason for an absence of an effect of urea upon refolding yield in buffer B is unclear. Solids fractions refolded in buffer 2 appeared (by eye) to be less turbid, than those refolded using buffer 1. This suggests, but does not confirm, that the levels of

aggregation are lower. Co-aggregation of protein with refolding target protein is theorized to be the reason refolding yields are reduced in the presence of contaminating proteins (Maachupalli-Reddy et al., 1997). Lower levels of aggregation of contaminating protein, mean that removal of these proteins by urea and sonication become less important. It is apparent that the effect of sonication is insignificant upon refolding yield, with both buffers, despite it being the most effective step in contaminant removal. This may suggest that the extra removal achieved by this by minimising particle size and maximising surface area for washing is not essential to providing higher refolding yields, and may compromise eventual yields.

Refolding pure trypsinogen with solids fractions from washes on XL10 solids fraction suggests that again Triton X appears to have the greatest effect on yield. It's position as the initial step in the washing scheme used or its significant effect on removal of lipids appears to improve yields.

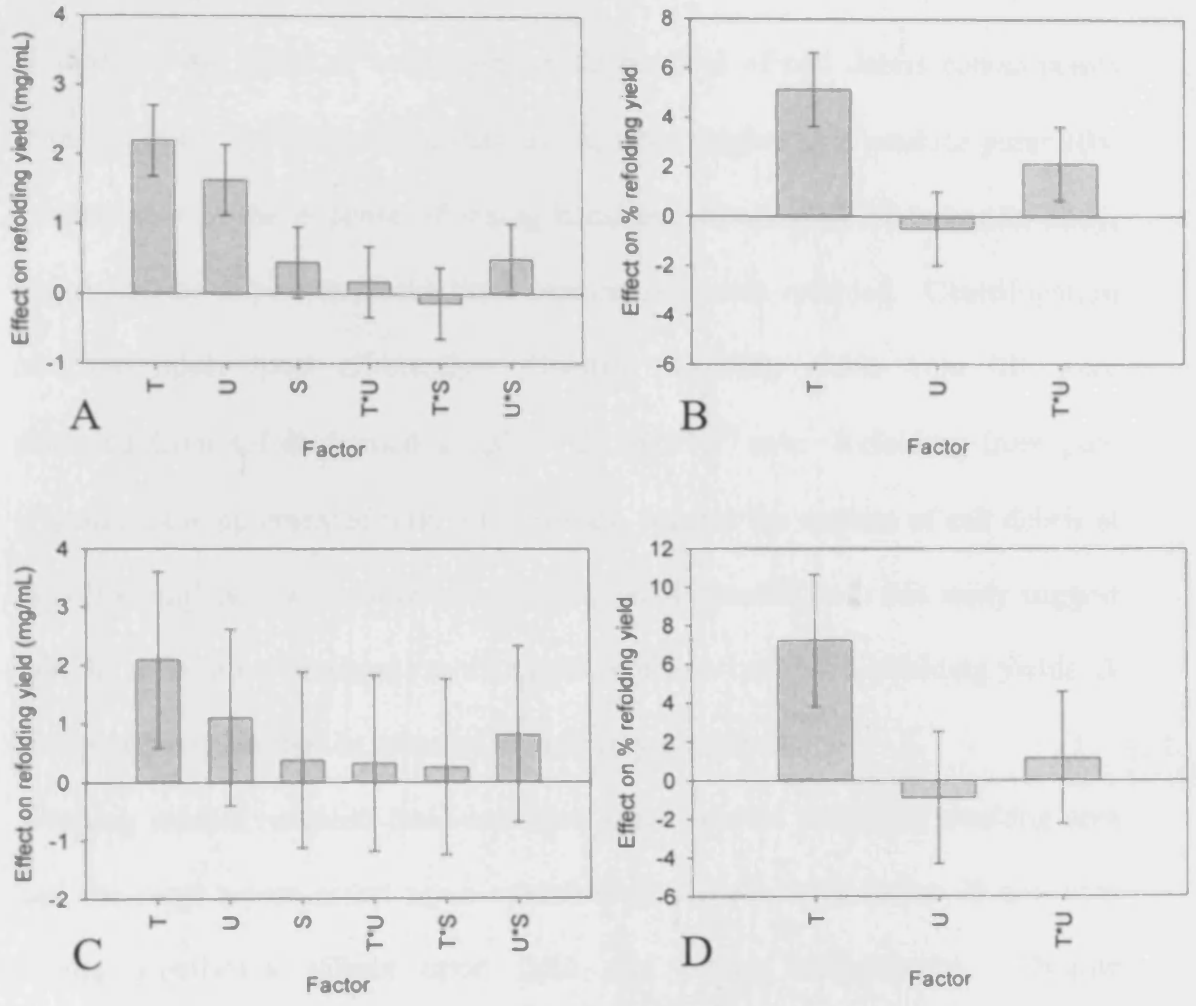


Figure 4.15 (A-D): Effect of washing factors upon refolding yield. In refolding buffer 1 from ELTRPI pellet (A) from XL10 pellet + 1mg/mL trypsinogen (B) and from refolding in buffer 2 from ELTRPI pellet (C) and XL10 + 1mg/mL trypsinogen (D).

4.4 Conclusions

Initial studies in this chapter that greatest yields confirmed are achieved with the buffer described in the literature (Buswell et al., 2002) at 4 °C. Factorial analysis revealed glucose as an additive had the most significant effect overall. A further factorial experiment revealed the significant effects of detergent, redox ratio, salt

concentration, choice of buffering species on refolding yields. Further experiment optimised salt concentration.

A study of the effect of centrifugation on removal of cell debris contaminants from inclusion bodies revealed that, as expected, higher Q/Σ produce purer IBs, but this was at the expense of losing considerable amounts of inclusion body, which will be important to the final amount of protein refolded. Centrifugation removed lipids most efficiently. Greatest refolding yields from IBs were achieved from solids formed at $Q/\Sigma = 23.08 \times 10^{-9}$ m/s. Refolding from pure trypsinogen supplemented with XL10 solids, suggest the amount of cell debris at the refolding stage was critical to refolding yield. Results from this study suggest that the ratio of contaminants to the target protein is critical to refolding yields. A compromise is reached in terms of IB recovery and purity.

Washing studies revealed that sonication which should maximise washing area had the most major effect upon contaminants levels, with Triton X and urea having significant effects upon lipid and protein respectively. Despite sonication's significant negative effect upon contaminant levels it had the least significant effect upon refolding yields. The greatest effect on yield was achieved by Triton X treatment followed by urea. Given the impact of Triton X upon lipid concentrations it appears that removal of lipid might improve yields, but it's position as the first step in the process may improve yields considerably by removing proteolytic enzymes. In addition urea washing had a significant effect on refolding perhaps because of its ability to solubilise contaminating protein shown to effect yields.

In conclusion it appears that to afford the greatest refolding yields, tryspinogen inclusion bodies are best separated at $Q/\Sigma = 23.08 \times 10^{-9}$ m/s to derive a balance

between recovery and purity, washing IBs with urea and Triton X to remove contaminating lipid and protein, which appear to negatively affect refolding yields.

This chapter concludes the experimental studies in this thesis. The final chapter brings together the observations from this study and looks to further work that may be done.

5 Conclusions and Further Work

5.1 Thesis Aims

The primary aim of this thesis was to study the effect of process parameters upon refolding yields. The study first identified the importance, interaction and effects of factors within the refolding process itself, and then studied the influence of upstream purification steps upon refolding efficiency.

5.1.1 Effect of Mixing

The first part of this thesis attempted to understand how and why mixing affects refolding. A novel twin-impeller system incorporating a mini-paddle impeller located in the vicinity of the injection point at which denatured protein was added, was used to increase the local levels of energy dissipation experienced by the injected material with the aim of improving refolding yields. Mixing only affected yields during and immediately after denatured protein addition. This suggests that when operating refolding reactors, mixing may only be required to achieve homogeneity. Mixing beyond this point appears to have a negligible effect on refolding, and evidence from later elements of this work suggest it might even reduce yields. Analysis of lysozyme refolding yield, under a variety of conditions, revealed that dispersive mixing affected the yield. The beneficial effect of the mini-paddle impeller in providing a source of localised energy dissipation was limited to conditions where the bulk impeller intensity was low. The effects appeared to become more significant when injection times were longer, because of increased exposure of the injected material to the effect of energy dissipation of the mini-impeller. The results suggest that for fed-batch

protein refolding systems, where mixing has been shown to be a critical factor, the local energy dissipation experienced in the vicinity of the injection point is critical to the refolding yields achieved. It is therefore thought that where refolding reactions are affected by mixing, reactor design should be such that denatured protein added sufficiently dispersed to maximise refolding yields.

5.1.2 Importance of Process Factors in Determining Refolding Yields

The second part of this study was aimed at understanding the relative importance of factors affecting refolding by dilution (e.g. mixing) and how they interact. Such knowledge is critical not only to maximising yield but also minimising the time to yield. The effect of five factors upon the dilution refolding of protein in a twin impeller fed-batch system using refold buffer containing only the oxidised form of the redox reagent was studied.

The five factors chosen were: bulk impeller Reynolds number, mini-impeller Reynolds number, injection rate of denatured protein, redox ratio, and guanidine hydrochloride (GdHCl) concentration. A 2^5 factorial experiment was conducted at an industrially relevant protein concentration using lysozyme as the test system. The study identified that in the system used, the guanidine hydrochloride concentration, redox ratio and injection rate were the most important factors in determining refolding yields. It was also demonstrated that under the conditions used mixing was not critical. For certain proteins, including lysozyme, sufficient levels of chaotrope in refold buffers can minimise the importance of sufficient mixing. This is thought to occur by reducing rate of aggregation such that is slower than mixing and is hence unaffected by its efficiency. Additional chaotrope does not always result in higher yields, so addition may not be suitable

in such systems and mixing may become important. Two interactions were seen to be important: redox ratio/guanidine hydrochloride concentration and guanidine hydrochloride concentration/ injection rate. Further analysis of the factorial data revealed that GdHCl concentration and injection rate were critical to determining the time to yield.

Conditions were also found at which high refolding yields could be achieved even with rapid injection and poor mixing efficiency. A comparative assessment was carried out with minimal mixing in a simple batch-refolding mode of operation. This revealed different behaviour to that of fed-batch. A graphical (windows of operation) analysis of the batch data suggested that optimal yields and productivity are obtained at high GdHCl concentrations (1.2 M) and redox ratios of unity or greater. This part of the thesis study emphasised the importance of studying interactions between factors to gain maximal yields in the shortest time possible.

5.1.3 The Impact of Downstream Processing on Refolding

The first two chapters of this thesis focussed on understanding the importance of dilution refolding parameters upon refolding and the effect of upstream steps upon the refolding reaction were not considered. Inefficiencies in IB purification steps mean that IBs solubilised for refolding are typically contaminated. These contaminants may increase proteolysis (Wong et al., 1996); reduce refolding yields (Georgiou and Valax, 1999; Maachupalli-Reddy et al., 1997) and/or impact on steps downstream of refolding (Thatcher et al., 1996). Studies were conducted to look at the effect of IB purification steps on contaminant removal and the subsequent effects on refolding yields, taking trypsinogen as the target protein.

Refolding of trypsinogen typically results in poor refolding yields (Buswell et al., 2002). Experiments showed that addition of glucose, NaCl/KCl, and adaptation of the redox ratio may improve the refolding yields generated.

Two IB processing steps were studied: centrifugation to separate IBs from cell debris and subsequent washing steps to remove IB contaminants. For each of these the levels of lipid, DNA and protein are measured and the refolding yield under each condition compared. Analysis of centrifugation processes revealed an expected compromise between IB purity and yield with centrifugal flow rate mimicked in the ultra-scaled-down system used. It appeared that amount of IB recovered appeared to be critical to the eventual refolding yield. The relative levels of target protein to contaminants appeared to be key to the yields generated; with a threshold level before protein refolding was affected. Centrifugation appeared to remove lipids more efficiently than any other contaminant.

Analysis of the effect of washing stages revealed that maximising washing area (achieved by sonication) was key to removing maximal amounts of contaminants, but washing with Triton X impacted on refolding yields most significantly.

5.1.4 Overall Conclusions

It is evident from this thesis that process condition must be carefully chosen to maximise refolding yields. The thesis has highlighted the importance of certain factors upon the refolding of a model protein, and the interactions between these factors. It is important to note that the relative importance of each factor can be incredibly protein specific, it is not the aim of these studies to suggest that these factors will affect the refolding of all proteins. It is however expected that if the factors revealed in this study are important they may interact as described. This

study revealed that the efficiency of IB processing steps can be critical to refolding yields. The degree of importance of these downstream processing steps is defined by the ease of the refolding step, with levels of contaminant being more critical for the refolding of some protein than others. Regardless of its effect on refolding it is advisable to ensure maximal efficiency in IB purification steps to reduce the burden on steps downstream of refolding.

5.2 Further Work

This thesis has studied the influence of parameters on refolding of factors both within the process of refolding itself and steps upstream in some detail. Areas of work that could be useful to study still remains.

It would be of interest to study the interaction between centrifugation and washing steps together, as this thesis only considered the impact of a single centrifugation condition upon washing. It would be beneficial to learn how any interactions between washing and centrifugation affect refolding yields and the levels of contaminants removed. By adjusting these conditions accordingly greater IB purity and refolding yields may be achieved.

Although this thesis has considered the effect of IB purification upon refolding yield, it has not considered the effect upon the efficiency of steps downstream of refolding, such as chromatographic polishing steps. In certain situations it is possible that contaminants levels may not be sufficient to effect refolding but may affect the efficiency of such chromatographic steps. It would be interesting to examine the tradeoffs between efficiency in IB purification, refolding and polishing steps. Such studies may indicate that certain purification steps could be

omitted without affecting the overall efficiency of refolding and post-refolding purification. Ultimately their removal should generate an increase in whole-process efficiency and economy.

6 References

References

Ahmad F and Bigelow CC. 1982. Estimation of the free energy of stabilization of ribonuclease A, lysozyme, α-lactalbumin, and myoglobin. J Biol Chem 257:12935-12938.

Ahmed AK, Schaffer SW, and Wetlaufer DB. 1975. Nonenzymic reactivation of reduced bovine pancreatic ribonuclease by air oxidation and by glutathione oxidoreduction buffers. J Biol Chem 250:8477-8482.

Ahn JH, Lee YP, and Rhee JS. 1997. Investigation of refolding condition for *Pseudomonas fluorescens* lipase by response surface methodology. J Biotechnol 54:151-160.

Al-Karadaghi S, Cedergren-Zeppenzauer E, and Houmoller S. 1994. Refined crystal structure of liver alcohol dehydrogenase-NADH complex at 1.8 A resolution. Acta Crystallogr D Biol Crystallogr 50:793-807.

Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD. 1994. Molecular biology of the cell. New York; London:Garland.

Altamarino MM, Garcia C, Possani LD, and Fersht AN. 1999. Oxidative refolding chromatography: folding of the scorpion toxin Cn5. Nat Biotechnol 17:187-191.

Amir D, Kraus S, and Haas E. 1992. Detection of local structures in reduced unfolded bovine pancreatic trypsin inhibitor. Proteins Struct Funct Genet 13:162-173.

Anfinsen CB. 1973. Principles that govern the folding of protein chains. Science 181:223-230.

Anfinsen CB and Haber E. 1961. Study on the reduction and reformation of protein disulfide bonds. J Biol Chem 236:1361-1363.

Arakawa T and Timasheff SN. 1982. Stabilisation of protein structure by sugars. Biochemistry 21:6536-6544.

Armstrong N, De Lencastre A, and Gouaux E. 1999. A new protein folding screen: Application to the ligand binding domains of glutamate and kainite receptor and to lysozyme and carbonic anhydrase. Protein Sci 8:1475-1483.

Assirelli M, Bujalski W, Eaglesham A, and Nienow AW. 2002. Study of micromixing in a stirred tank using a Rushton turbine, comparison of feed positions and other mixing devices. Trans IChemE 80:855-863.

Baldwin R,Creighton T. 1980. Gesellschaft für Biologische Chemie Konferenz (28th : 1979 : Regensburg): Protein folding : proceedings of the 28th

Conference of the German Biochemical Society, held at the University of Regensburg, Regensburg, West Germany, September 10-12, 1979. Amsterdam:Elsevier. 217 p.

Ballew R, Salbeko J, and Gruebele M. 1996. Observation of distinct nanosecond and microsecond protein folding events. Nat Struct Biol 3:923-926.

Bam NB, Cleland JL, and Randolph T. 1996. Molten globule intermediate of recombinant human growth hormone: stabilization with surfactants. Biotechnol Prog 12:801-809.

Batas B, Schiraldi C, and Chandhuri JB. 1999. Inclusion body purification and protein refolding using microfiltration and size exclusion chromatography. J Biotechnol 68:149-155.

Bessette PH, Aslund F, Beckwith J, and Georgiou G. 1999. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. Proc Natl Acad Sci U S A 96:13703-13708.

Betts SD. 1998. Inclusion body formation in *E. coli* observed by electron microscopy. http://web.mit.edu/king-lab/www/research/Scott/Scott-Research.html

Beynon RJ,Easterby JS. 1996. Buffer Solutions. Oxford; New York:IRL Press at Oxford University Press.

Bhat KM, Sumner IG, Perry BN, Pickersgill RW, and Goodenough P. 1991. A novel method for the purification of porcine phospholipase A2 expressed in E. coli. Biophys Res Comm 176:371-377.

Bird RE, Hardman KD, Jacobson JW, Johnson S, Kaufman BM, Lee SM, Lee T, Pope SH, Riordan GS, and Whitlow M. 1988. Single-chain antigen-binding proteins. Science 242:423-426.

Bourne JR. 1985. Mixing in single-phase chemical reactors. In: Harnby N, Edwards MF, Nienow AW, editors. Mixing in the process industries. London. Butterworths. p 184-199.

Bowden GA and Georgiou G. 1990. Folding and aggregation of beta-lactamase in the periplasmic space of *Escherichia coli*. J Biol Chem 265:16760-16765.

Bowden GA, Petrides D, and Georgiou G. 1991. Structure and morphology of protein inclusion bodies in *Escherichia coli*. Bio/Technology 9:730-

Boychyn M, Yim SSS, Bulmer M, More J, Bracewell DG, and Hoare M. 2004. Performance prediction of industrial centrifuges using scale down models. Bioprocess Biosyst Eng 26:385-391.

Branden C,Tooze J. 1999. Introduction to protein structure. New York:Garland.

Brandts JF, Halvorson HR, and Brennan M. 1975. Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. Biochemistry 14:4953-4963.

Brygelson JD, Onuchic JN, Socci ND, and Wolynes PG. 1995. Funnels, pathways, and the energy landscape of protein folding: a synthesis. Prot Struct Funct Genet 21:167-195.

Buchner J and Rudolph R. 1991. Renaturation, purification and characterization of recombinant of recombinant Fab-fragments produced *Escherichia coli*. Bio/Technology 9:157-162.

Burley S K and Petsko GA. 1985. Aromatic-aromatic interaction: A mechanism of protein structure stabilization. Science 229:23-38.

Buswell AM, Ebtinger M, Vertés AA, and Middelberg APJ. 2002. Effect of operating variables on the yield of recombinant trypsinogen for a pulse-fed dilution-refolding reactor. Biotechnol Bioeng 77:435-444.

Buswell AM and Middelberg APJ. 2002. Critical analysis of lysozyme refolding kinetics. Biotechnol Prog 18:470-475.

Buswell AM and Middelberg APJ. 2003. A new kinetic scheme for lysozyme refolding and aggregation. Biotechnol Bioeng 83:567-577.

Campbell MK. 1999. Biochemistry. Philadelphia:Saunders College Publishing : Harcourt Brace College Publishers.

Carlson JD and Yarmush ML. 1992. Antibody assisted protein refolding. Bio/Technology 10:86-91.

Carrió MM, Cubarsi R, and Villaverde A. 2000. Fine architecture of bacterial inclusion bodies. FEBS Lett 471:7-11.

Carrió MM and Villaverde A. 2001. Protein aggregation as bacterial inclusion bodies is reversible. FEBS Lett 489:29-33.

Carrió MM and Villaverde A. 2002. Construction and deconstruction of bacterial inclusion bodies. J Biotechnol 96:3-12.

Carrió MM and Villaverde A. 2003. Role of molecular chaperones in inclusion body formation. FEBS Lett 537:215-221.

Chen W, Helenius J, Braakman I, and Helenius A. 1995. Cotranslational folding and calnexin binding during glycoprotein synthesis. Proc Natl Acad Sci U S A 92:6229-6233.

Cheng YSE, Kwoh D, Kwoh TJ, Soltveldt B, and Zipser D. 1981. Stabilization of a degradable protein by its overexpression in *Escherichia coli*. Gene 14:121-130.

Cho TH, Ahm SJ, and Lee EK. 2002. Refolding of protein inclusion bodies directly from E.coli homogenate using expanded bed adsorption chromatography. Bioseparation 10:189-196.

Choi J, Jeong K, Kim S, and Lee SY. 2000. Efficient secretory production of alkaline phosphatase by high cell density culture of recombinant *Escherichia coli* using the *Bacillus sp.* endoxylanase signal sequence. Appl Microbiol Biotechnol 53:640-645.

Chothia C. 1976. The nature of the accessible and buried surfaces in proteins. J Mol Biol 105:1-14.

Cleland JL, Builder SE, Swartz JR, Winkler M, Chang JY, and Wang DIC. 1992a. Polyethylene glycol enhanced protein refolding. Bio/Technology 10:1013-1019.

Cleland JL, Hedgepeth C, and Wang DIC. 1992b. Polyethylene glycol enhanced refolding of bovine carbonic anhydrase B. Reaction stoichiometry and refolding model. J Biol Chem 267:13327-13334.

Cleland JL and Randolph T. 1992. Mechanism of polyethylene glycol interaction with the molten globule folding intermediate of bovine carbonic anhydrase B. J Biol Chem 267:3147-3153.

Cleland JL and Wang DIC. 1990a. Cosolvent assisted protein refolding. Bio/Technology 8:1274-1278.

Cleland JL and Wang DIC. 1990b. Refolding and aggregation of bovine carbonic anhydrase B: quasi-elastic light scattering analysis. Biochemistry 29:11072-11078.

Cleland JL, Wang DIC. 1993. *In-vitro* protein refolding. In: Stephanspoulos G, editors. Vol.3 Bioprocessing. Weinheim, Germany. VCH. p 528-558.

Cohen SN, Chang ACY, Boyer HW, and Helling RB. 1973. Construction of biologically functional bacterial plasmids *in vitro*. Proc Natl Acad Sci U S A 70:3240-3244.

Creighton TE. 1986. Folding of proteins adsorbed reversibly to ion-exchange resins. In: Oxender DL, editors. UCLA symposia on molecular and cell biology, New series. New York. Alan R Liss.

Datar RV, Cartwright T, and Rosen CG. 1993. Process economics of animal cell and bacterial fermentations: a case study analysis of tissue plasminogen activator. Bio/Technology 11:349-357.

Davis JM, Narachi MA, Alton NK, and Arakawa T. 1987. Structure of human tumor necrosis factor alpha derived from recombinant DNA. Biochemistry 26:1322-1326.

De-Bernardez Clark E. 1998. Refolding of recombinant proteins. Curr Opin Biotechnol 9:157-163.

De-Bernardez Clark E. 2001. Protein refolding for industrial processes. Curr Opin Biotechnol 12:202-207.

De-Bernardez Clark E, Hevehan D, Szela S, and Maachupalli-Reddy J. 1998. Oxidative renaturation of hen egg–white lysozyme. Folding vs aggregation. Biotechnol Prog 14:47-54.

Derman A, Prinz W, Belin D, and Beckwith J. 1993. Mutations that allow disulfide bond formation in the cytoplasm of *Escherichia coli*. Science 262:1744-1747.

Diamond R. 1974. Real-space refinement of the structure of hen egg-white lysozyme. J Mol Biol 82:371-391.

Dill KA. 1990. Dominant forces in protein refolding. Biochemistry 29:7133-7155.

Dobson CM, Sali A, and Karplus M. 1998. Protein folding: A perspective from theory and experiment . Angew Chem Int Ed 37:868-893.

Dong X-Y, Huang Y, and Sun Y. 2004. Refolding kinetics of denatured-reduced lysozyme in the presence of folding aids. J Biotechnol 114:135-142.

Dong X-Y, Yang H, and Sun Y. 2000. Lysozyme refolding with immobilized GroEL column chromatography. J Chromatogr A 878:197-204.

Doran PM. 1995. Bioprocess engineering principles. London:Academic Press Limited.

Dorin G, McAlary P, and Wong K. 1996. Bacterial production of hydrophobic polypeptides. WO1996US09155 19960605:

Ellis RJ. 2001. Macromolecular crowding: obvious but underappreciated. Trends Biochem Sci 26:597-604.

Ellis RJ and Hartl FU. 1999. Principles of protein folding in the cellular environment. Curr Opin Struct Biol 9:102-110.

Ellis R. 1997. Molecular chaperones: avoiding the crowd. Curr Biol 7:R531-R533.

Epstein CJ and Goldberger RF. 1963. A study of the factors influencing the reactivation of reduced egg white lysozyme. J Biol Chem 238:1380-1383.

Fahnert B, Lilie H, and Neubauer P. 2004. Inclusion bodies: Formation and utilisation. Adv Biochem Engin/Biotechnol 89:93-142.

Falconer RJ, O'Neill BK, and Middelberg APJ. 1998. Chemical treatment of *Escherichia coli*. II. Direct extraction of recombinant protein from cytoplasmic inclusion bodies in intact cells. Biotechnol Bioeng 57:381-386.

Falconer RJ, O'Neill BK, and Middelberg APJ. 1999. Chemical treatment of *Escherichia coli*: 3. Selective extraction of a recombinant protein from cytoplasmic inclusion bodies in intact Cells. Biotechnol Bioeng 62:455-460.

Fang JZ and Lee DJ. 2001. Micromixing efficiency in static mixer. Chem Eng Sci 56:3797-3802.

Fischer B, Sumner I, and Goodenough P. 1993. Isolation, renaturation and formation of disulfide bonds of eukaryotic proteins expressed in *Escherichia coli* as inclusion bodies. Biotechnol Bioeng 41:3-13.

Fournier MC, Falk L, and Villermaux J. 1996. A new parallel competing reaction system for assessing micromixing efficiency-determination of micromixing time by a simple mixing model. Chem Eng Sci 51:5187-5192.

Fromage N, Denefle P, Cambou B, Duchese M, Joyeux C, Kovarik S, Marin J, Imbault F, and Uzan A. 1991. Synthesis, purification and biological properties of a truncated mutant form of human tissue plasminogen activator produced in *E. Coli*. Fibrinolysis 5:187-190.

Frydman J, Erdjument-Bromage H, Tempst P, and Hartl F. 1999. Co-translational domain folding as the structural basis for the rapid *de novo* folding of firefly luciferase. Nat Struct Mol Biol 6:697-705.

Garcia P, Desmadril M, Minard P, and Yon J. 1995. Evidence for residual structures in the unfolded form of yeast phosphoglycerate kinase. Biochemistry 34:397-404.

Gardiner, S A M. 1988. Process design for recombinant protein recovery from inclusion bodies.PhD thesis. University College London.

Geli V, Baty D, Knibiehler M, Lloubes R, Pessegue B, Shire D, and Lazdunski C. 1989. Synthesis and sequence-specific proteolysis of a hybrid protein (colicin A::growth hormone releasing factor) produced in Escherichia coli. Gene 80:129-136.

Geng X, Bai Q, Zhang Y, Li X, and Wu D. 2004. Refolding and purification of interferon-gamma in industry by hydrophobic interaction chromatography. J Biotechnol 113:137-149.

Georgiou G and Valax P. 1999. Isolating inclusion bodies from bacteria. Methods in Enzymology 309:48-58.

Ghelis C,Yon J. 1982. Protein folding. New York:Academic Press.

Gilmore GN. 1977. A modern approach to comprehensive chemistry. Cheltenham:Stanley Thornes.

Glynou K, Ioannou PC, and Christopoulos T. 2003. One-step purification and refolding of recombinant photoprotein aequorin by immobilized metal-ion affinity chromatography. Protein Expression Purif 27:384-390.

Goldberg ME, Rudolph R, and Jaenicke R. 1991. A kinetic study of the competition between renaturation and aggregation during the refolding of denatured-reduced egg-white lysozyme. Biochemistry 30:2790-2797.

Guichardon P and Falk L. 2000. Characterisation of micromixing efficiency iodide-iodate reaction system. Part I: experimental procedure. Chem Eng Sci 55:4233-4243.

Guillard F and Trägårdh C. 2003. Mixing in industrial Rushton turbine-agitated reactors under aerated conditions. Chem Eng Process 42:373-386.

Guo L. 2001. Simultaneous purification and renaturation of recombinant human interferon-alpha expressed by E. coli by high-performance hydrophobic interaction chromatography. Chin J Chromatogr 19:301-303.

Halenbeck R, Kawasaki E, Wrin J, and Koths K. 1989. Renaturation and purification of biologically active recombinant human macrophage colony-stimulating factor expressed in *E. Coli*. Bio/Technology 7:710-715.

Hames B, Hooper N, Houghton J. 1997. Instant Notes In Biochemistry. Oxford, UK:BIOS Sciences Publishers.

Hart RA, Rinas U, and Bailey JE. 1990. Protein composition of Vitreoscilla hemoglobin inclusion bodies produced in *Escherichia coli*. J Biol Chem 265:12728-12733.

Hevehan D and De-Bernardez Clark E. 1997. Oxidative renaturation of lysozyme at high concentrations. Biotechnol Bioeng 54:221-230.

Hibbert, E G. 2004. Engineering and molecular biology approaches to improving trypsin-based bioprocesses. PhD thesis. UCL. London, UK.

Hoare M and Dunnill P. 1989. Biochemical engineering challenges of purifying useful proteins. Phil Trans R Soc London B 324:497-507.

Hohenblum H, Vorauer-Uhl K, Katinger H, and Mattanovich D. 2004. Bacterial expression and refolding of human trypsinogen. J Biotechnol 109:3-11.

Hoogendoorn CJ and den Hartog AP. 1967. Model studies on mixers in the viscous flow region. Chem Eng Sci 22:1689-1699.

Hsu Y-R and Arakawa T. 1985. Structural studies on acid unfolding and refolding of recombinant human interferon gamma. Biochemistry 24:7959-7963.

Izard J and Limberger R. 2003. Rapid screening method for quantitation of bacterial cell lipids from whole cells. J Microbiol Methods 55:411-418.

Jackson DA, Symons RH, and Berg P. 1972. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: Circular SV40 DNA molecules containing Lambda Phage genes and galactose operon of *Escherichia coli*. Proc Natl Acad Sci U S A 69:2904-2909.

Jagannadham MV and Balasubramanian D. 1985. The molten globule intermediate form in the folding pathway of human carbonic anhydrase. FEBS Lett 188:326-330.

Jin K, Thomas O, and Dunhill P. 1994. Monitoring recombinant inclusion body recovery in an industrial disc stack centrifuge. Biotechnol Bioeng 43:455-460.

Jürgen B, Lin HY, Riemschneider S, Scharf C, Neubauer P, Schmid R, Hecker M, and Schweder T. 2000. Monitoring of genes that respond to overproduction of an insoluble recombinant protein in Escherichia coli glucose-limited fed-batch fermentations. Biotechnol Bioeng 70:217-224.

Katoh S and Katoh Y. 2000. Continuous refolding of lysozyme with fed-batch addition of denatured protein solution. Process Biochem (Oxford, UK) 35:1119-1124.

Katoh S, Sezai Y, Yamaguchi T, Katoh Y, Yagi H, and Nohara D. 1999. Refolding of enzymes in a fed-batch operation. Process Biochem (Oxford, UK) 35:297-300.

Kauzmann W. 1959. Some factors in the interpretation of protein denaturation. Adv Protein Chem 14:1-64.

Kawaguchi Y, Shimizu N, Nishimori K, Uozumi T, and Beppu T. 1984. Renaturation and activation of calf prochymosin produced in an insoluble form in *Escherichia coli*. J Biotechnol 1:307-315.

Kiefhaber T. 1995. Kinetic traps in lysozyme refolding. Proc Natl Acad Sci U S A 92:9029-9033.

Kim PS and Baldwin RL. 1982. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. Annu Rev Biochem 51:459-489.

Kirby AJ. 2001. The lysozyme mechanism sorted- after 50 years. Nat Struct Biol 8:737-739.

Kuharski RA and Rossky PJ. 1984. Solvation of hydrophobic species in aqueous urea solution: a molecular dynamics study. J Am Chem Soc 106:5794-5800.

Kuwajima K. 1989. The molten globule state as a clue for understanding the folding and cooperativity of globular protein structure. Prot Struct Genet 6:87-103.

Kuwajima K. 1996. The molten globule state of alpha-lactalbumin. FASEB J 10:102-109.

Langley KE, Berg TF, Strickland TW, Fenton DM, Boone TC, and Wypych J. 1987. Recombinant-DNA-derived bovine growth hormone from Escherichia coli. 1. Demonstration that the hormone is expressed in reduced form, and isolation of the hormone in oxidized, native form. Eur J Biochem 163:313-321.

Lee CT, Buswell AM, and Middelberg APJ. 2002. The influence of mixing on lysozyme renaturation during refolding in an oscillatory flow and stirred tank reactor. Chem Eng Sci 57:1679-1684.

Lee CT, Morreale G, and Middelberg APJ. 2004. Combined in-fermenter extraction and cross-flow microfiltration for improved inclusion body processing. Biotechnol Bioeng 85:103-113.

Lemercier G, Baklara N, and Santarelli X. 2003. On-column refolding of an insoluble histidine tag recombinant exopolyphosphatase from Trypanosoma brucei overexpressed in Escherichia coli. J Chromaogr B 786:305-309.

Levinthal C. 1968. Are there pathways for protein folding? J Chim Phys Phys Chim Bio 65:44-45.

Levitt M. 1981. Effect of proline residues on protein folding. J Mol Biol 145:251-263.

Li M, Zhang GF, and Su Z. 2002. Dual gradient ion-exchange chromatography improved refolding yield of lysozyme. J Chromatogr A 959:113-120.

Li M, Zhi-Guo S, and Janson J-C. 2004. *In vitro* protein refolding by chromatographic procedures. Protein Expression Purif 33:1-10.

Lin L-N and Brandts J. 1978. Further Evidence Suggesting That the Slow Phase in Protein Unfolding and Refolding is Due to Proline Isomerisation: A Kinetic Study of Carp Parvalbumins. Biochemistry 17:4102-4110.

Lin WW and Lee DJ. 1997. Micromixing effects in aerated stirred tanks. Chem Eng Sci 52:3837-3842.

Lobban PE and Kaiser AD. 1973. Enzymatic end-to-end joining of DNA molecules. J Mol Biol 78:453-471.

London J, Skrzynia C, and Goldberg ME. 1974. Renaturation of *Escherichia coli* tryptophanase after exposure to 8 M urea. Evidence for the existence of nucleation centers. Eur J Biochem 47:409-415.

Maa Y and Hsu CC. 1997. Protein denaturation by combined effect of shear and air-liquid interface. Biotechnol Bioeng 54:503-512.

Maachupalli-Reddy J, Kelley BD, and De-Bernardez Clark E. 1997. Effect of inclusion body contaminants on the oxidative renaturation of hen egg white lysozyme. Biotechnol Prog 13:144-150.

Machida S, Ogawa S, Xiaohua S, Takahara T, Fujii K, and Hayashi K. 2000. Cycloamylose as an efficient artificial chaperone for protein refolding. FEBS Lett 486:131-135.

Maeda Y, Yamuda H, Ueda T, and Imoto T. 1996. Effect of additives on the renaturation of reduced lysozyme in the presence of 4 M urea. Protein Sci 9:461-465.

Mannall GJ, Titchener Hooker NJ, and Dalby PA. 2006. A critical assessment of the impact of mixing on dilution refolding. Biotechnol Bioeng 93:955-963.

Marston FAO, Lowe PA, Shoemacher JM, White S, and Angal S. 1984. Purification of calf prochymosin (prorennin) synthesized in *Escherichia coli*. Bio/Technology 2:800-804.

Matouschek A, Kellis JT, Serrano L, Bycroft M, and Fersht AN. 1990. Transient folding intermediates characterized by protein engineering. Nature 346:440-445.

Mattes R. 2006. The production of improved Tissue-Type Plasminogen Activator in *Escherichia coli*. Semin Thromb Hemost 27:325-335.

Mertz JE and Davis RW. 1972. Cleavage of DNA by R 1 restriction endonuclease generates cohesive end. Proc Natl Acad Sci U S A 69:3370-3374.

Middelberg APJ. 2002. Preparative protein refolding. Trends Biotechnol 20:437-443.

Miranker A, Robinson CV, Radford SE, Aplin RT, and Dobson CM. 1993. Detection of transient protein folding populations by mass spectrometry. Science 262:896-900.

Mitraki A and King J. 1989. Protein folding intermediates and inclusion body formation. Bio/Technology 7:690-697.

Narayana SVL and Argos P. 1984. Structural patterns in globular proteins. Int J Pept Protein Res 24:25-39.

Némethy G and Scheraga HA. 1962a. Structure of water and hydrophobic bonding in proteins. I. A model for the thermodynamic properties of liquid water. J Chem Phys 36:3382-3400.

Némethy G and Scheraga HA. 1962b. Structure of water and hydrophobic bonding in proteins. II. Model for the thermodynamic properties of aqueous solutions of hydrocarbons. J Chem Phys 36:3401-3417.

Oberg K, Chrunyk BA, Wetzel R, and Fink A. 1994. Native-like secondary structure in interleukin-1.beta. inclusion bodies by attenuated total reflectance FTIR. Biochemistry 33:2628-2634.

Oldshue JY. 1983. Fluid Mixing Technology. New York:McGraw Hill.

Onuchic JN, Nyemeyer H, Garcia AE, Chahine J, and Socci ND. 2000. The energy landscape theory of protein folding: insights into folding mechanisms and scenarios. Adv Protein Chem 53:87-152.

Onuchic JN, Wolynes PG, Luthney-Schultzen Z, and Socci ND. 1995. Toward an outline of the topography of a realistic protein-folding funnel. Proc Natl Acad Sci U S A 92:3626-3630.

Orsini G and Goldberg ME. 1978. The renaturation of reduced chymotrypsinogen A in guanidine HCl. Refolding versus aggregation. J Biol Chem 253:3453-3458.

Pain R. 1987. Protein folding for pleasure and for profit. Trends Biochem Sci 12:309-312.

Pauling L and Corey RB. 1951. The pleated sheet, a new layer configuration of polypeptide chains. Proc Natl Acad Sci U S A 37:251-256.

Pauling L, Corey RB, and Branson HR. 1951. The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. Proc Natl Acad Sci U S A 37:205-211.

Peters TJ. 1975. The Plasma Proteins: Structure, Function and Genetic Control. In: Putnam Fv1, editors. New York; London. Academic Press. p 147-

Petrides D, Sapidou E, and Calandranis J. 1995. Computer-aided process analysis and economic-evaluation for biosynthetic insulin production- a case study. Biotechnol Bioeng 48:529-541.

Porter DH and Cardenas JM. 1980. Analysis of the renaturation kinetics of bovine muscle pyruvate kinase. Biochemistry 19:3447-3452.

Proba K, Ge L, and Pluckthun A. 1995. Functional antibody single-chain fragments from the cytoplasm of Escherichia coli: influence of thioredoxin reductase (TrxB). Gene 159:203-207.

Przybycien TM, Dunn J, Valax P, and Georgiou G. 1994. Secondary structure characterization of beta-lactamase inclusion bodies. Protein Eng 7:131-136.

Ptitsyn OB, Pain R, Semisotnov G, Zervovnik E, and Razgulyaev OI. 1990. Evidence for a molten globule state as a general intermediate in protein folding. FEBS Lett 26:21-24.

Ptitsyn OB and Rashin AA. 1973. Stagewise mechanism of protein folding. Doklady Akademii Nauk SSSR 213:473-475.

Radford SE, Dobson CM, and Evans PA. 1992. The folding of hen lysozyme involves partially structured intermediates and multiple pathways. Nature 358:302-307.

Rehm BHA, Qi Q, Beermann B, Hinz HJ, and Steinbuchel A. 2001. Matrix-assisted in vitro refolding of Pseudomonas aeruginosa class II polyhydroxyalkanoate synthase from inclusion bodies produced in recombinant Escherichia coli. Biochem J 358:263-268.

Renzeperis D, Jonsson T, and Sauer RT. 1999. Acceleration of the refolding of Arc repressor by nucleic acids and other polyanions. Nat Struct Mol Biol 6:569-573.

Rinas U and Bailey JE. 1992. Protein compositional analysis of inclusion bodies produced in recombinant Escherichia coli. Appl Microbiol Biotechnol 59:609-614.

Rinas U and Bailey JE. 1993. Overexpression of bacterial hemoglobin causes incorporation of pre-beta-lactamase into cytoplasmic inclusion bodies. Appl Environ Microbiol 59:561-566.

Rinas U, Boone TC, and Bailey JE. 1993. Characterization of inclusion bodies in recombinant Escherichia coli producing high levels of porcine somatotropin. J Biotechnol 28:313-320.

Rose GD, Geselowitz AR, Lesser GJ, Lee RH, and Zehfus MH. 1985. Hydrophobicity of amino acid residues in globular proteins. Science 229:834-838.

Roux P, Delepierre M, Goldberg ME, and Caffotte A-F. 1997. Kinetics of secondary structure recovery during the refolding of reduced hen egg white lysozyme. J Biol Chem 272:24843-24849.

Rozema D and Gellman S. 1996. Artificial chaperone-assisted refolding of carbonic anhydrase B. J Biol Chem 271:3478-3487.

Rudolph R. 1996. Successful protein refolding on an industrial scale. In: Cleland JL, Craick C, editors. Protein engineering: Principles and practice. New York. Wiley.

Rudolph R, Böhm G, Lilie H, Jaenicke R. 1997. Folding proteins. In: Creighton T, editors. Protein Function a Practical Approach. Oxford. Oxford University Press. p 57-99.

Rudolph R, Engelhard M, and Jaenicke R. 1976. Kinetics of refolding and reactivation of Rabbit-Muscle Aldolase after acid dissociation. Eur J Biochem 67:455-462.

Rudolph R and Fischer S. 1990. Process for obtaining renatured proteins. 298274:

Rudolph R, Heider I, and Jaenicke R. 1977a. Effect of coenzymes and temperature on the process of in vitro refolding and reassociation of lactic dehydrogenase isoenzymes. Biochemistry 16:5527-5531.

Rudolph R, Heider I, Westhof E, and Jaenicke R. 1977b. Mechanism of refolding and reactivation of lactic dehydrogenase from pig heart after dissociation in various solvent media. Biochemistry 16:3384-3390.

Rudolph R, Zettlmeissl G, and Jaenicke R. 1979. Reconstitution of lactic dehydrogenase. Noncovalent aggregation vs. reactivation. 2. Reactivation of irreversibly denatured aggregates. Biochemistry 18:5572-5575.

Ruiyin C and Krishna MA. 2001. Method for recovery of proteins prepared by recombinant DNA procedures.

Russell CS and Clarke LA. 1999. Recombinant proteins for genetic disease. Clinical Genetics 55:389-394.

Sakono M, Kawashima Y, Ichinose H, Maruyama T, Kamiya N, and Goto M. 2004. Direct refolding of inclusion bodies using reversed micelles. Biotechnol Prog 20:1783-1787.

Sawano H, Koumoto Y, Ohta K, Sasaki Y, Segawa S-I, and Tachibana H. 1992. Efficient in vitro folding of the three-disulfide derivatives of hen lysozyme in the presence of glycerol. FEBS Lett 303:11-14.

Schein CH and Noteborn MHM. 1988. Formation of soluble recombinant proteins in *Escherichia Coli* is favored by lower growth temperature. Bio/Technology 6:291-294.

Shiue SJ and Wong CW. 1984. Studies on homogenization efficiency of various agitators in liquid blending. Can J Chem Eng 62:602-609.

Smith A, Santama N, Dacey S, Edwards M, Bray RC, Thorneley RN, and Burke JF. 1990. Expression of a synthetic gene for horseradish peroxidase C in Escherichia coli and folding and activation of the recombinant enzyme with Ca2+ and heme. J Biol Chem 265:13335-13343.

Sohl JL, Jaswal SS, and Agard DA. 1998. Unfolded conformations of α-lytic protease are more stable than its native state. Nature 395:817-819.

Sophianopoulos AJ, Rhodes CK, Holcomb DN, and Van Kimmenade A. 1962. Physical studies of lysozyme. I. Characterisation. J Biol Chem 237:1107-1112.

Speed MA, Wang DIC, and King J. 1996. Specific aggregation of partially folded polypeptide chains: The molecular basis of inclusion body composition. Nat Biotechnol 14:1283-1287.

Stempfer G, Holl-Neugebauer R, and Rudolph R. 1996. Improved refolding of an immobilized fusion protein. Nat Biotechnol

Storrs SB, Przybycien TM. 1991. Commercial-scale refolding of recombinant methionyl bovine somatotropin. In: Georgiou G, De-Bernandez Clark E, editors. Protein Refolding. Washington, DC. American Chemical Society. p 197-205.

Stryer L. 1995. Biochemistry. New York:W.H. Freeman and Company.

Sugimoto S and Yokoo Y. 1991. Isolation and characterization of recombinant eel growth hormone expressed in *Escherichia coli*. Biotechnol Lett 13:289-394.

Sugimoto S, Yokoo Y, Inni Y, and Hirano T. 1991. Large-scale purification and refolding of recombinant eel growth hormone. Agric Biol Chem 55:1635-1637.

Sundari CS, Raman B, and Balasubramanian D. 1999. Artificial chaperoning of insulin, human carbonic anhydrase and hen egg lysozyme using linear dextrin chains – a sweet route to the native state of globular proteins. FEBS Lett 443:215-219.

Tandon S and Horowitz PM. 1987. Detergent-assisted refolding of guanidinium chloride-denatured rhodanese. The effects of the concentration and type of detergent. J Biol Chem 262:4486-4491.

Tanford C. 1968. Protein denaturation. Adv Protein Chem 23:121-282.

Thatcher DR. 1990. Recovery of therapeutic proteins from inclusion bodies: problems and process strategies. Biochem Soc Trans 18:234-235.

Thatcher DR, Wilks P, Chaudhuri J. 1996. Inclusion bodies and refolding. In: Price NC, editors. Proteins labfax. Oxford. Bios Scientific/ Academic Press. p 119-130.

Titchener Hooker NJ, Gritsis D, Olbrich R, Mannweiler K, Gardiner SAM, Fish NM, and Hoare M. 1991. Integrated process design for producing and recovering proteins from inclusion bodies. Pharmaceutical Technology International 1991:42-48.

Tokatlidis K, Dhurjati D, Millet J, Beguin P, and Aubert JP. 1991. High activity of inclusion bodies formed in *Escherichia coli* overproducing *Clostridium thermocellum* endoglucanase D. FEBS Lett 282:205-208.

Tsong TY. 1982. Viscosity-dependent conformational relaxation of ribonuclease A in the thermal unfolding zone. Biochemistry 21:1493-1497.

Tsumoto K, Ejima D, Kumagi I, and Arakawa T. 2003. Practical considerations in refolding proteins from inclusion bodies. Protein Expression Purif 28:1-8.

Tsumoto K, Shinoki K, Kondo H, Uchikawa M, Juji T, and Kumagi I. 1998. Highly efficient recovery of functional single-chain Fv fragments from inclusion bodies overexpressed in Escherichia coli by controlled

introduction of oxidizing reagents- applications to a human single-chain Fv fragment. J Immunol Methods 219:119-129.

Ujvari A, Aron R, Eisenhaure T, Cheng E, Parag HA, Smicun Y, Halaban R, and Hebert DN. 2001. Translation rate of human tyrosinase determines its N-linked glycosylation level. J Biol Chem 276:5924-5931.

Uversky VN and Ptitsyn OB. 1994. Partly folded" state, a new equilibrium state of protein molecules: four-state guanidinium chloride-induced unfolding of beta-lactamase at low temperature. Biochemistry 33:2782-2791.

Uversky VN and Ptitsyn OB. 1996. Further evidence on the equilibrium "pre-molten globule state": four-state guanidinium chloride-induced unfolding of carbonic anhydrase B at low temperature. J Mol Biol 255:215-228.

Valax P, Georgiou G. 1991. Folding and aggregation of RTEM beta-lactamase. Protein Refolding. Washington DC. American Chemical Society. p 97-109.

Valax P and Georgiou G. 1993. Molecular characterization of β-lactamase inclusion bodies produced in *Escherichia Coli.* 1.Composition. Biotechnol Prog 9:539-547.

Van Kimmenade A, Bond MW, Schumacher JH, Laquoi C, and Kastelein RA. 1988. Expression, renaturation and purification of recombinant human interleukin 4 from *Escherichia Coli.* Eur J Biochem 173:109-114.

Varnein JP, Smith T, Rosenblum CI, Vongs A, Murphy BA, Nunes C, Mellin TN, King JJ, Burgess BW, Junker B, Chou M, Hey P, Frazier E, MacIntyre DE, Van der Ploeg LH, and Tota MR. 1998. Production of leptin in *Escherichia coli*: a comparison of methods. Curr Opin Biotechnol 12:202-207.

Wang C, Castro AF, Wilkes D, and Altenberg G. 1999. Expression and purification of the first nucleotide-binding domain and linker region of human multidrug resistance gene product: comparison of fusions to glutathione S-transferase, thioredoxin and maltose-binding protein. Biochem J 338:77-81.

Werner MH, Clore GM, Gronenborn AM, Kondoh A, and Fischer RJ. 1994. Refolding proteins by gel filtration chromatography. FEBS Lett 345:125-130.

West SM, Chaudhuri JB, and Binding P. 1998. Improved protein refolding using hollow-fibre membrane dialysis. Biotechnol Bioeng 57:590-599.

Wetlaufer DB and Ristow S. 1973. Acquisition of the three dimensional structure of proteins. Annu Rev Biochem 42:135-158.

Wetlaufer DB and Xie Y. 1995. Control of aggregation in protein refolding: A variety of surfactants promote renaturation of carbonic anhydrase II. Protein Sci 4:1535-1543.

White FHJ. 1961. Regeneration of native secondary and tertiary structures by air oxidation of reduced ribonuclease. J Biol Chem 236:1353-1360.

Wildegger G and Kiefhaber T. 1997. Three-state model for lysozyme folding: Triangular folding mechanism with an energetically trapped intermediate. J Mol Biol 270:294-304.

Wilhelm OG, Jaskunas SR, Vlahos CJ, and Bang NU. 1990. Functional properties of the recombinant kringle-2 domain of tissue plasminogen activator produced in *Escherichia coli*. J Biol Chem 265:14606-14611.

Willis MS, Hogan JK, Prabhakar P, Liu X, Tsai K, Wei Y, and Fox T. 2005. Investigation of protein refolding using a fractional factorial screen: A study of reagent effects and interactions. Protein Sci 14:1818-1826.

Winkler ME, Brigman T, and Marks BJ. 1986. The purification of fully active recombinant transforming growth factor α produced in *Escherichia coli*. J Biol Chem 261:13838-13843.

Wolynes PG, Onuchic JN, and Thirumalai D. 1995. Navigating the folding routes. Science 267:1618-1620.

Wong HH, O'Neill BK, and Middelberg APJ. 1996. Centrifugal recovery and dissolution of recombinant Gly-IGF-II inclusion-bodies: the impact of feed-rate and re-centrifugation on protein yield. Bioseparation 6:185-192.

Wong HH, O'Neill BK, and Middelberg APJ. 1997. Centrifugal processing of cell debris and inclusion bodies from recombinant *Escherichia coli*. Bioseparation 6:361-372.

Woodley JM and Titchener Hooker NJ. 1996. The use of windows of operation as a bioprocess design tool. Bioprocess Eng 14:263-268.

Wu H. 1931. Studies on denaturation of proteins. A theory of denaturation. Adv Protein Chem 46:6-26.

Yasuda M, Murakami Y, Sowa A, Ogino H, and Ishikawa H. 1998. Effects of additives on refolding of a denatured protein. Biotechnol Prog 14:601-606.

Yon JM. 2002. Protein folding in the post-genomic era. J Cell Mol Med 6:307-327.

Yoshimoto M and Kuboi R. 1999. Oxidative refolding of Denatured/Reduced lysozyme utilizing the chaperone-like function of liposomes and immobilized liposome chromatography. Biotechnol Prog 15:480-487.

Zettlmeissl G, Rudolph R, and Jaenicke R. 1979. Reconstitution of lactic dehydrogenase. Noncovalent aggregation vs. reactivation. 1. Physical properties and kinetics of aggregation. Biochemistry 18:5567-5571.

Appendix 1- Suppliers

Beckman-Coulter (UK), London Road, High Wycombe, Buckinghamshire, HP11 1JU

Biorad (UK), Maxted Road, Hemel Hempstead, Hertfordshire,HP2 7DX.

Eli-Lilly, Indianopolis, Indiana, USA.

Eppendorf, Chivers Way, Histon, Cambridgeshire CB4 9ZR.

Fisher Scientific (UK),Bishop Meadow Road, Loughborough, Leicestershire, LE11 5RG

GE Healthcare, GE-75 184, Uppsala, Sweden.

Grant instruments, Shepreth, Cambridgeshire, SG8 6GB.

Heidolph instruments Walpersdorfers st., 91126 Schwabach, Germany.

IKA, Janke & Kunkel st., D-79219, Staufen, Germany.

Invitrogen, 3 Fountain Drive, Inchinnan Business Park Paisley, UK

Jencons Scientific, Stanbridge Road, Leighton Buzzard, Bedfordshire, LU7 4UA

MSE, London

Oxoid Ltd, Wade Road, Basingstoke, Hampshire, RG24 8PW

Phenomenex, Hurdsfield Industrial estate, Macclesfield, Cheshire, SK10 2BN.

Sigma-Aldrich (UK), Fancy Road, Poole, Dorset, BH12 4QH

Tecan, Seestrasse, Männedorf, Switzerland.

Thermo Spectronic, Linden Avenue, Rochester, NY, USA.

UVP, 11th Street, Upland, California, USA.

VWR International, Hunter Boulevard, Magna Park, Lutterworth, Leicestershire, LE17 4XN

Worthington Biochemical, Lorne Laboratories Ltd, Ruscombe Lane, Twyford, Reading, Berkshire, RG10 9NJ.